GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 12:10:38 ; Search time 2162 Seconds
(without alignments)
8969.086 Million cell updates/sec

Title: US-10-049-568-1
Perfect score: 474
Sequence: 1 gcccagatttattcagtggc.....ggagatgccacctgagttaa 474

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	474	100.0	474	6	AX088165	Sequence
2	439.2	92.7	1191	6	AX128507	Sequence
3	439.2	92.7	2274	6	AX549337	Sequence
4	439.2	92.7	2274	9	AF190500	Homo sapi
5	439.2	92.7	2467	6	BD135243	Novel mam
6	439.2	92.7	3584	6	BD135246	Novel mam
7	222	46.8	187431	9	AC107219	Homo sapi
8	171.8	36.2	1065	6	BD141178	Novel G p
9	171.8	36.2	1068	6	AX148176	Sequence
10	171.8	36.2	1473	6	BD141176	Novel G p
11	171.8	36.2	1545	6	BD141175	Novel G p
12	171.8	36.2	1830	6	BD141172	Novel G p
13	171.8	36.2	2049	6	AX385037	Sequence
14	171.8	36.2	2142	6	AX451566	Sequence
15	171.8	36.2	2181	6	AX709190	Sequence
16	171.8	36.2	2190	6	BD141181	Novel G p
17	171.8	36.2	2193	6	AX385032	Sequence
18	171.8	36.2	2214	6	AX451562	Sequence
19	171.8	36.2	2262	6	BD141180	Novel G p
20	171.8	36.2	2265	6	AX385027	Sequence
21	171.8	36.2	2436	9	AF453828	Homo sapi
22	171.8	36.2	2838	9	AF403384	Homo sapi
23	170.6	36.0	220994	2	AC128964	Rattus no
24	170.6	36.0	230365	2	AC121415	Rattus no
c 25	170.6	36.0	243409	2	AC098607	Rattus no
c 26	163	34.4	158394	2	AC122462	Mus muscu
c 27	163	34.4	253348	2	AC132143	Mus muscu
28	156.8	33.1	321	6	AX384516	Sequence
29	146.2	30.8	2214	6	AX385045	Sequence
30	146.2	30.8	2539	10	AF346501	Mus muscu
c 31	106.8	22.5	1018	6	AX147820	Sequence
c 32	106.8	22.5	1018	6	AX521869	Sequence
33	106.8	22.5	44585	6	AX646897	Sequence
34	106.8	22.5	170522	9	AL138708	Human DNA
c 35	92.4	19.5	180673	10	AC068627	Mus muscu
c 36	92.4	19.5	213462	10	AC077689	Mus muscu
c 37	86	18.1	211211	2	AC121029	Rattus no
38	86	18.1	239190	2	AC098990	Rattus no
c 39	73	15.4	167899	2	BX005476	Danio rer
c 40	73	15.4	195380	5	AL929075	Zebrafish
41	61.4	13.0	176565	2	AL136106	Homo sapi
42	57.2	12.1	653	4	AY196483	Equus cab
43	48.8	10.3	1015	6	AX254453	Sequence
44	47	9.9	104992	2	AC005504	Plasmodiu
45	47	9.9	169546	2	AC004157	Plasmodiu

ALIGNMENTS

RESULT 1
AX088165
LOCUS AX088165 474 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 1 from Patent WO0114548.
ACCESSION AX088165
VERSION AX088165.1 GI:13397078
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Duecker,K.
TITLE New g-protein coupled receptor and dna sequences thereof
JOURNAL Patent: WO 0114548-A 1 01-MAR-2001;

```
            MERCK PATENT GmbH (DE)
FEATURES             Location/Qualifiers
     source          1..474
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             <1..474
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC34686.1"
                     /db_xref="GI:13397079"
                     /translation="AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQ
                     VKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIGYSAINSALN
                     PILYTLTTRPFKEMIHRFWHNYRQRKSMDSKGIRKHMLHHSSGGKCGHCRRCHLS"
BASE COUNT      144 a     92 c      88 g     150 t
ORIGIN

  Query Match             100.0%;  Score 474;  DB 6;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 1e-99;
  Matches  474;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474


RESULT 2
AX128507
LOCUS       AX128507                1191 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 15 from Patent WO0131014.
ACCESSION   AX128507
VERSION     AX128507.1  GI:14134974
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Vogeli,G., Wood,L.S. and Merchant,K.
  TITLE     G protein-coupled receptors expressed in brain
  JOURNAL   Patent: WO 0131014-A 15 03-MAY-2001;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1..1191
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..1191
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC38938.1"
                     /db_xref="GI:14134975"
                     /translation="MFRPLVNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENLLAS
                     IIQRVFVWVVSAVTCFGNIFVICMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGG
                     FDLKFRGEYNKHAQLWMESTHCQLVGSLAILSTEVSVLLLTFLTLEKYICIVYPFRCV
                     RPGKCRTITVLILIWITGFIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIY
                     SVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTD
                     ALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF
                     WYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPELMKPDLFTYPCEMSLISQSTRL
                     NSYS"
BASE COUNT      340 a    229 c     226 g     396 t
ORIGIN

  Query Match             92.7%;  Score 439.2;  DB 6;  Length 1191;
  Best Local Similarity   96.8%;  Pred. No. 1.1e-91;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 702

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 762

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 822

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        823 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 882

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |    ||||||||||||||
Db        883 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 942

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 1002

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db       1003 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 1060

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1061 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 1114


RESULT 3
AX549337
LOCUS       AX549337                2274 bp    DNA     linear   PAT 26-NOV-2002
DEFINITION  Sequence 622 from Patent WO02061087.
ACCESSION   AX549337
VERSION     AX549337.1  GI:25813979
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burmer,G.C., Roush,C.L. and Brown,J.P.
  TITLE     Antigenic peptides, such as for G protein-coupled receptors
            (GPCRs), antibodies thereto, and systems for identifying such
            antigenic peptides
  JOURNAL   Patent: WO 02061087-A 622 08-AUG-2002;
            Lifespan Biosciences, Inc. (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .2274
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT     669 a      460 c      421 g      724 t
ORIGIN

  Query Match           92.7%;  Score 439.2;  DB 6;  Length 2274;
  Best Local Similarity  96.8%;  Pred. No. 1e-91;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 1785

Qy         61 GTTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 GTTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 1845

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 1905

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTGTGAGTGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db       1906 ACTGATGCATTATGCTGGATACCCATTTTGTGAGTGAAATTCTTTCACTGCTTCAGGTA 1965

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1966 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 2025

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 2085

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAACATATGCTC 420
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2086 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAACATATGCTC 2143

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2144 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2197

RESULT 4
AP190500
LOCUS       AP190500                2274 bp    mRNA    linear   PRI 10-AUG-2001
DEFINITION  Homo sapiens leucine-rich repeat-containing G protein-coupled
            receptor 7 (LGR7) mRNA, complete cds.
ACCESSION   AP190500
VERSION     AP190500.1  GI:10441729
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2274)
  AUTHORS   Hsu,S.Y., Kudo,M., Chen,T., Nakabayashi,K., Bhalla,A., van der
            Spek,P.J., van Duin,M. and Hsueh,A.J.
  TITLE     The three subfamilies of leucine-rich repeat-containing G
            protein-coupled receptors (LGR): identification of LGR6 and LGR7
            and the signaling mechanism for LGR7
  JOURNAL   Mol. Endocrinol. 14 (8), 1257-1271 (2000)
  MEDLINE   20388592
   PUBMED   10935549
REFERENCE   2  (bases 1 to 2274)
  AUTHORS   Hsu,S.Y. and Hsueh,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-1999) GYN/OB, Stanford University, MSOB S385,
            Stanford, CA 94305, USA
FEATURES             Location/Qualifiers
     source          1. .2274
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .2274
                     /gene="LGR7"
     CDS             1. .2274
                     /gene="LGR7"
                     /codon_start=1
                     /product="leucine-rich repeat-containing G protein-coupled
                     receptor 7"
                     /protein_id="AAG17167.1"
                     /db_xref="GI:10441730"
                     /translation="MTSGSVYFILIFGKEYFSHGGGQDVKCSLGYFPCGNITKCLPQL
                     LHCNGVDDCGNQADEDNCGDNNGWSMQFDKYFASYYKMTSQYPPEATPECVLVGSVPV
                     QCLCQGLELDCDETNLRAVPSVSSNVTAMSLQWNLIRKLPPDCPKNYIHDLQKLYLQNN
                     KITSISIYAFRGLNSLTKLYLSHNRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTF
                     YGLNSLILVLNNNVLTRLPDKPSLHWLDLESGNHIHNLRNLTFISCSNLTVL
                     VMRKNKINSLNENTFAPLQKLDELDLGSNKIRNLPPLIFKDLKKLSQLNLSYNPIQKI
                     QANQFDYLVKLKSLSLDGIRISNIQQRMFRPLNNLSHIIYFKFPQYCGYAPHVRSCKPN
                     TGCISSLSNLLAEIIQRVFVWVSAVTCPGNIFVICMRPYIRSENKLYAMSIISLCCA
                     DCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLVGSLAILSFKVSVLLLTPLTL
                     EKYICIVYPFRCVRPGKCRTITVLLLIWITGFIVAPIPLSNKEFFKNYYGTNGVCFPL
                     HSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFTSVHQSAITATEIRNQVKKEM
                     ILAKRFFIVFTDALCWIPIFVVKFLSLLQVKIPGTITSWVVIFILPINSALNPILYT
                     LTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMMPLQEMPPEKLMKPDLPTY
                     PCEMSLISQSTRLNSYS"
BASE COUNT     669 a      460 c      421 g      724 t
ORIGIN

  Query Match           92.7%;  Score 439.2;  DB 9;  Length 2274;
  Best Local Similarity  96.8%;  Pred. No. 1e-91;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 1785

Qy         61 GTTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 GTTTTTTCCTATGGAAGCATGTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 1845

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 1905

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTGTAGCGAACCCTCTTTCATCGCTTCAGGTA 240
              |||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db       1906 ACTGATGCATTATGCTGGATACCCATTTTGTGAGTGAAATTCTTTCACTGCTTCAGGTA 1965

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1966 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 2025

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 2085

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAACATATGCTC 420
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2086 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAACATATGCTC 2143

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2144 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2197

RESULT 5
BD135245
LOCUS       BD135245                2467 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Novel mammalian G protein-coupled receptor having extracellular
            leucine-rich repeating domain.
ACCESSION   BD135245
VERSION     BD135245.1  GI:23230190
```

```
KEYWORDS    JP 2002507406-A/3.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2467)
  AUTHORS   Hsueh,A.J.W., Hsu,S.Y., Liang,S.G. and Spe,P.J.V.D.
  TITLE     Novel mammalian G protein-coupled receptor having extracellular
            leucine-rich repeating domain
  JOURNAL   Patent: JP 2002507406-A 3 12-MAR-2002;
            THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY,
            AKZO NOBEL NV
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002507406-A/3
            PD   12-MAR-2002
            PF   25-MAR-1999 JP 2000537903
            PR   26-MAR-1998 US     60/079501
            PI   AARON J W HSUEH,SHEAU YU HSU,SHAN GUANG LIANG,PETRUS JOHANNES
            PI   VAN DER SPEK
            PC   C12N15/09,A01K67/027,C07K14/705,C07K16/28,C12N1/15,C12N1/19,
            PC   C12N1/21,
            PC   C12N5/10,C12P21/08,C12N15/00,C12N5/00
            CC   Novel mammalian G protein-coupled receptor having  CC
              extracellular
            CC   leucine-rich repeating domain
            FH   Key            Location/Qualifiers
            FT   source         1. .2467
            FT                  /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .2467
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      747 a     487 c     474 g     759 t
ORIGIN

  Query Match            92.7%;  Score 439.2;  DB 6;  Length 2467;
  Best Local Similarity  96.8%;  Pred. No. 1e-91;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 1978

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 2038

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 2098

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db       2099 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 2158

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |    ||||||||||||||
Db       2159 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 2218

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 2278

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db       2279 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 2336

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       2337 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2390
```

```
RESULT 6
BD135246
LOCUS       BD135246                3584 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Novel mammalian G protein-coupled receptor having extracellular
            leucine-rich repeating domain.
ACCESSION   BD135246
VERSION     BD135246.1  GI:23230191
KEYWORDS    JP 2002507406-A/4.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3584)
  AUTHORS   Hsueh,A.J.W., Hsu,S.Y., Liang,S.G. and Spe,P.J.V.D.
  TITLE     Novel mammalian G protein-coupled receptor having extracellular
            leucine-rich repeating domain
  JOURNAL   Patent: JP 2002507406-A 4 12-MAR-2002;
            THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY,
            AKZO NOBEL NV
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002507406-A/4
            PD   12-MAR-2002
            PF   25-MAR-1999 JP 2000537903
            PR   26-MAR-1998 US     60/079501
            PI   AARON J W HSUEH,SHEAU YU HSU,SHAN GUANG LIANG,PETRUS JOHANNES
            PI   VAN DER SPEK
            PC   C12N15/09,A01K67/027,C07K14/705,C07K16/28,C12N1/15,C12N1/19,
            PC   C12N1/21,
            PC   C12N5/10,C12P21/08,C12N15/00,C12N5/00
            CC   Novel mammalian G protein-coupled receptor having  CC
              extracellular
            CC   leucine-rich repeating domain
            FH   Key            Location/Qualifiers
            FT   source         1. .3584
            FT                  /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .3584
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT     1124 a     670 c     647 g    1139 t      4 others
ORIGIN

  Query Match            92.7%;  Score 439.2;  DB 6;  Length 3584;
  Best Local Similarity  96.8%;  Pred. No. 9.8e-92;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1744 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 1803

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1804 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 1863

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1864 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 1923

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db       1924 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 1983

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |    ||||||||||||||
Db       1984 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 2043

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2044 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 2103

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||  |||||||||||||||||
```

```
Db      2104 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 2161

Qy       421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      2162 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2215


RESULT 7
AC107219
LOCUS       AC107219               187431 bp    DNA     linear   PRI 27-MAR-2002
DEFINITION  Homo sapiens BAC clone RP11-575B4 from 4, complete sequence.
ACCESSION   AC107219
VERSION     AC107219.5  GI:19526153
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 187431)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
    PUBMED  9847074
REFERENCE   2  (bases 1 to 187431)
  AUTHORS   Isak,A., Meyer,R. and Boyer,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-575B4
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 187431)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 187431)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 187431)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 187431)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Mar 26, 2002 this sequence version replaced gi:18855160.
            ---------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            ---------- Summary Statistics
                Center project name: H_NH0575B04
            ----------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is CTD-2005A22; the clone sequenced
            to the right is RP11-663M18.  Actual start of this clone is at base
            position 1 of RP11-575B4; actual end is at base position 187431 of
            RP11-575B4.

            Sequence derived from one plasmid subclone, base position 102011 to
            102054.

FEATURES             Location/Qualifiers
     source          1..187431
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /map="4"
                     /clone="RP11-575B4"
                     /clone_lib="RPCI-11"
     repeat_region   1..408
                     /rpt_family="L1"
     repeat_region   563..697
                     /rpt_family="Alu"
     repeat_region   829..938
                     /rpt_family="Alu"
     repeat_region   1227..1326
                     /rpt_family="MIR"
     repeat_region   1327..2602
                     /rpt_family="L1"
     repeat_region   2973..3009
                     /rpt_family="AT_rich"
     repeat_region   3233..3534
                     /rpt_family="Alu"
     repeat_region   3582..3616
                     /rpt_family="AT_rich"
     repeat_region   3685..3772
                     /rpt_family="L2"
     repeat_region   4141..4297
                     /rpt_family="MaLR"
     repeat_region   4298..4593
                     /rpt_family="Alu"
     repeat_region   4594..4804
                     /rpt_family="MaLR"
     repeat_region   5020..5436
                     /rpt_family="MaLR"
     repeat_region   5483..5506
                     /rpt_family="AT_rich"
     repeat_region   6006..6170
                     /rpt_family="MaLR"
     repeat_region   6709..6748
                     /rpt_family="L2"
     repeat_region   6863..7446
                     /rpt_family="L1"
     repeat_region   7454..8078
                     /rpt_family="ERV1"
     repeat_region   8079..8175
                     /rpt_family="L1"
     repeat_region   8176..8477
```

```
                    /rpt_family="Alu"
     misc_feature   8176..8189
                    /note="match to EST AI683385 (NID:g4893567) tx66e09.x1"
     repeat_region  8478..8745
                    /rpt_family="L1"
     repeat_region  8746..9052
                    /rpt_family="Alu"
     repeat_region  9053..9124
                    /rpt_family="L1"
     repeat_region  9126..9148
                    /rpt_family="(TTTTC)n"
     misc_feature   9681..9926
                    /note="similar to Mus musculus EST BB297224
                    (NID:g16401965)"
     repeat_region  9921..10022
                    /rpt_family="L1"
     misc_feature   10225..10665
                    /note="similar to Mus musculus EST BB297224
                    (NID:g16401965)"
     misc_feature   10870..11388
                    /note="match to EST AI653269 (NID:g4737248) wb44e10.x1"
     misc_feature   11115..11388
                    /note="match to EST AI652121 (NID:g4736100) wb47c02.x1"
     repeat_region  12295..12442
                    /rpt_family="Alu"
     repeat_region  12443..12734
                    /rpt_family="Alu"
     repeat_region  12735..12900
                    /rpt_family="Alu"
     repeat_region  13540..13819
                    /rpt_family="Alu"
     repeat_region  14177..14344
                    /rpt_family="Alu"
     repeat_region  14358..15112
                    /rpt_family="L1"
     repeat_region  15130..15158
                    /rpt_family="(TA)n"
     repeat_region  15194..15487
                    /rpt_family="Alu"
     repeat_region  15575..15853
                    /rpt_family="Alu"
     misc_feature   16606..16891
                    /note="match to EST AI375172 (NID:g4175162) tc10b12.x1"
     misc_feature   16789..17233
                    /note="match to EST AI805323 (NID:g5391889) te32h05.x1"
     repeat_region  17004..17094
                    /rpt_family="L2"
     misc_feature   complement(17219..17220)
                    /note="match to EST AA122078 (NID:g1678116) zk93g10.r1"
     misc_feature   17228..17513
                    /note="match to EST AA122078 (NID:g1678116) zk93g10.r1"
     repeat_region  17508..17621
                    /rpt_family="L1"
     misc_feature   17610..18078
                    /note="match to EST AI089319 (NID:g3428378) qb16d05.x1"
     repeat_region  17655..17710
                    /rpt_family="L1"
     misc_feature   17849..18074
                    /note="match to EST AA626188 (NID:g2538575) zv89a01.s1"
     repeat_region  18230..18295
                    /rpt_family="Alu"
     repeat_region  18298..18349
                    /rpt_family="Alu"

  Query Match             46.8%;  Score 222;  DB 9;  Length 187431;
  Best Local Similarity   97.8%;  Pred. No. 1.5e-41;
  Matches  225;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         29 TTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTCCTATGGAAGCATGTTTTATA 88
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13203 TAGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTCCTATGGAAGCATGTTTTATA 13262

Qy         89 GTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGA 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13263 GTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGA 13322

Qy        149 TCCTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTT 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      13323 TCCTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTT 13382

Qy        209 TTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATA 258
              ||||| |||| ||||||||||||||||||||||||||||||||||| |||
Db      13383 TTGTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGTACAATA 13432


RESULT 8
BD141178
LOCUS       BD141178               1065 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Novel G protein-coupled receptor protein and its DNA.
ACCESSION   BD141178
VERSION     BD141178.1  GI:23236123
KEYWORDS    WO 0204640-A/7.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1065)
  AUTHORS   Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.
  TITLE     Novel G protein-coupled receptor protein and its DNA
  JOURNAL   Patent: WO 0204640-A 7 17-JAN-2002;
            TAKEDA CHEMICAL INDUSTRIES LTD,TAKEO MORIYA,TAKASHI ITO, YASUSHI
            SHINTANI, NOBUYUKI MIYAJIMA
COMMENT     OS   Homo sapiens (human)
            PN   WO 0204640-A/7
            PD   17-JAN-2002
            PF   06-JUL-2001 WO 2001JP005878
            PR   07-JUL-2000 JP 00P   211989,18-DEC-2000 JP 00P   383794 PI
            TAKEO MORIYA,TAKASHI ITO,YASUSHI SHINTANI,NOBUYUKI MIYAJIMA PC
            C12N15/12,C12N1/21,C07K14/705,C07K16/28,C12P21/02,C12Q1/68, PC
            A61K38/00,
            PC   A61K45/00,A61K48/00,A61P1/00,A61P3/00,A61P9/00,A61P25/28, PC
            A61P29/00,
            PC   A61P35/00,A61P37/00,G01N33/15,G01N33/50,G01N33/53,G01N33/566//
            PC   (C12N1/21,C12R1:19),(C12P21/02,C12R1:19)
            CC   Novel G protein-coupled receptor protein and its DNA. FH    Key
                      Location/Qualifiers
            FT   source          1..1065
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1..1065
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      277 a     214 c     213 g     361 t
ORIGIN

  Query Match             36.2%;  Score 171.8;  DB 6;  Length 1065;
  Best Local Similarity   67.4%;  Pred. No. 1.2e-29;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTCC 69
              |||||  | | |||||| || ||| | || |||     || ||| ||||||| || ||||||
Db        568 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 627

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
              |||   |  ||||| | |     |||| |||  |||  |     | || ||| || |||||
Db        628 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 687

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCA 189
                 ||  || |||| ||    |||| || ||||| || |||||||| ||  ||||||
Db        688 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTATAGTGTTCTCTGATGCC 747

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
               | |||||||| ||   | |||||||  |||    ||||| ||    | | ||| ||||||||
Db        748 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 807
```

```
Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
              |  || || || || ||| |||||||  ||      | |  |||||||||||||||| |||
Db        808 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 867

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
              || ||||||||||| || ||  |       ||||| ||  |  |  |  | | |   ||| ||
Db        868 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 926


RESULT 9
AX148176
LOCUS       AX148176                1068 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 17 from Patent WO0136471.
ACCESSION   AX148176
VERSION     AX148176.1  GI:14347081
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chen,R., Dang,H.T. and Lowitz,K.P.
  TITLE     Endogenous and non-endogenous versions of human g protein-coupled
            receptors
  JOURNAL   Patent: WO 0136471-A 17 25-MAY-2001;
            Arena Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1068
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      278 a    215 c    214 g    361 t
ORIGIN

  Query Match             36.2%;  Score 171.8;  DB 6;  Length 1068;
  Best Local Similarity   67.4%;  Pred. No. 1.2e-29;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
              |||||  | | |||||| || ||| | || |||     || ||| ||||||| || ||||||
Db        568 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 627

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
              |||     |  |||||  | |        ||||  |||  |||  |      | || ||| || |||||
Db        628 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 687

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCA 189
                   ||  || |||| ||      |||| || ||||| || |||||||| ||  |||||||
Db        688 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 747

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
               | |||||||| ||   | |||||||  |||     ||||| ||   |  | ||| |||||||||
Db        748 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 807

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
              |  || || || || ||| |||||||  ||      | |  |||||||||||||||| |||
Db        808 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 867

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
              || ||||||||||| || ||  |       ||||| ||  |  |  |  | | |   ||| ||
Db        868 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 926


RESULT 10
BD141176
LOCUS       BD141176                1473 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Novel G protein-coupled receptor protein and its DNA.
ACCESSION   BD141176
VERSION     BD141176.1  GI:23236121
KEYWORDS    WO 0204640-A/5.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1473)
  AUTHORS   Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.
  TITLE     Novel G protein-coupled receptor protein and its DNA
  JOURNAL   Patent: WO 0204640-A 5 17-JAN-2002;
            TAKEDA CHEMICAL INDUSTRIES LTD,TAKEO MORIYA,TAKASHI ITO, YASUSHI
            SHINTANI, NOBUYUKI MIYAJIMA
COMMENT     OS   Homo sapiens (human)
            PN   WO 0204640-A/5
            PD   17-JAN-2002
            PF   06-JUL-2001 WO 2001JP005878
            PR   07-JUL-2000 JP 00P   211989,18-DEC-2000 JP 00P   383794 PI
            TAKEO MORIYA,TAKASHI ITO,YASUSHI SHINTANI,NOBUYUKI MIYAJIMA PC
            C12N15/12,C12N1/21,C07K14/705,C07K16/28,C12P21/02,C12Q1/68, PC
            A61K38/00,
            PC   A61K45/00,A61K48/00,A61P1/00,A61P3/00,A61P9/00,A61P25/28, PC
            A61P29/00,
            PC   A61P35/00,A61P37/00,G01N33/15,G01N33/50,G01N33/53,G01N33/566//
            PC   (C12N1/21,C12R1:19),(C12P21/02,C12R1:19)
            CC   Novel G protein-coupled receptor protein and its DNA. FH    Key
                      Location/Qualifiers
            FT   source          1..1473
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1..1473
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      405 a    310 c    278 g    480 t
ORIGIN

  Query Match             36.2%;  Score 171.8;  DB 6;  Length 1473;
  Best Local Similarity   67.4%;  Pred. No. 1.2e-29;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
              |||||  | | |||||| || ||| | || |||     || ||| ||||||| || ||||||
Db        976 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 1035

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
              |||     |  |||||  | |        ||||  |||  |||  |      | || ||| || |||||
Db       1036 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 1095

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCA 189
                   ||  || |||| ||      |||| || ||||| || |||||||| ||  |||||||
Db       1096 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 1155

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
               | |||||||| ||   | |||||||  |||     ||||| ||   |  | ||| |||||||||
Db       1156 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 1215

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
              |  || || || || ||| |||||||  ||      | |  |||||||||||||||| |||
Db       1216 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 1275

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
              || ||||||||||| || ||  |       ||||| ||  |  |  |  | | |   ||| ||
Db       1276 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 1334


RESULT 11
BD141175
LOCUS       BD141175                1545 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Novel G protein-coupled receptor protein and its DNA.
ACCESSION   BD141175
VERSION     BD141175.1  GI:23236120
KEYWORDS    WO 0204640-A/4.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1545)
  AUTHORS   Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.
```

```
TITLE      Novel G protein-coupled receptor protein and its DNA
JOURNAL    Patent: WO 0204640-A 4 17-JAN-2002;
           TAKEDA CHEMICAL INDUSTRIES LTD,TAKEO MORIYA,TAKASHI ITO, YASUSHI
           SHINTANI, NOBUYUKI MIYAJIMA
COMMENT    OS   Homo sapiens (human)
           PN   WO 0204640-A/4
           PD   17-JAN-2002
           PF   06-JUL-2001 WO 2001JP005878
           PR   07-JUL-2000 JP 00P   211989,18-DEC-2000 JP 00P   383794 PI
           TAKEO MORIYA,TAKASHI ITO,YASUSHI SHINTANI,NOBUYUKI MIYAJIMA PC
           C12N15/12,C12N1/21,C07K14/705,C07K16/28,C12P21/02,C12Q1/68, PC
           A61K38/00,
           PC   A61K45/00,A61K48/00,A61P1/00,A61P3/00,A61P9/00,A61P25/28, PC
           A61P29/00,
           PC   A61P35/00,A61P37/00,G01N33/15,G01N33/50,G01N33/53,G01N33/566//
           PC   (C12N1/21,C12R1:19),(C12P21/02,C12R1:19)
           CC   Novel G protein-coupled receptor protein and its DNA. FH   Key
                    Location/Qualifiers
           FT   source          1. .1545
           FT                   /organism='Homo sapiens (human)'.
FEATURES            Location/Qualifiers
     source         1. .1545
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
BASE COUNT      429 a    320 c    290 g    506 t
ORIGIN

  Query Match             36.2%;  Score 171.8;  DB 6;  Length 1545;
  Best Local Similarity   67.4%;  Pred. No. 1.2e-29;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
Db       1048 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 1107

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
Db       1108 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 1167

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCA 189
Db       1168 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 1227

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
Db       1228 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 1287

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
Db       1288 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 1347

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
Db       1348 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 1406

RESULT 12
BD141172
LOCUS       BD141172                1830 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Novel G protein-coupled receptor protein and its DNA.
ACCESSION   BD141172
VERSION     BD141172.1  GI:23236117
KEYWORDS    WO 0204640-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1830)
  AUTHORS   Moriya,T., Ito,T., Shintani,Y. and Miyajima,N.
  TITLE     Novel G protein-coupled receptor protein and its DNA
  JOURNAL   Patent: WO 0204640-A 1 17-JAN-2002;
            TAKEDA CHEMICAL INDUSTRIES LTD,TAKEO MORIYA,TAKASHI ITO, YASUSHI
```

```
            SHINTANI, NOBUYUKI MIYAJIMA
COMMENT     OS   Homo sapiens (human)
            PN   WO 0204640-A/1
            PD   17-JAN-2002
            PF   06-JUL-2001 WO 2001JP005878
            PR   07-JUL-2000 JP 00P   211989,18-DEC-2000 JP 00P   383794 PI
            TAKEO MORIYA,TAKASHI ITO,YASUSHI SHINTANI,NOBUYUKI MIYAJIMA PC
            C12N15/12,C12N1/21,C07K14/705,C07K16/28,C12P21/02,C12Q1/68, PC
            A61K38/00,
            PC   A61K45/00,A61K48/00,A61P1/00,A61P3/00,A61P9/00,A61P25/28, PC
            A61P29/00,
            PC   A61P35/00,A61P37/00,G01N33/15,G01N33/50,G01N33/53,G01N33/566//
            PC   (C12N1/21,C12R1:19),(C12P21/02,C12R1:19)
            CC   Novel G protein-coupled receptor protein and its DNA. FH   Key
                     Location/Qualifiers
            FT   source          1. .1830
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1830
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      516 a    377 c    351 g    586 t
ORIGIN

  Query Match             36.2%;  Score 171.8;  DB 6;  Length 1830;
  Best Local Similarity   67.4%;  Pred. No. 1.1e-29;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
Db       1333 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 1392

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
Db       1393 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 1452

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTTACTGATGCA 189
Db       1453 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 1512

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
Db       1513 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 1572

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
Db       1573 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 1632

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
Db       1633 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 1691

RESULT 13
AX385037
LOCUS       AX385037                2049 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 11 from Patent WO0214489.
ACCESSION   AX385037
VERSION     AX385037.1  GI:19578160
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Paszty,C.J., Gong,J., Daugherty,B. and Rogers,N.
  TITLE     Leucine-rich repeat-containing g-protein coupled receptor-8
            molecules and uses thereof
  JOURNAL   Patent: WO 0214489-A 11 21-FEB-2002;
            Amgen, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2049
```

```
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
     CDS            1..2049
                    /note="unnamed protein product"
                    /codon_start=1
                    /protein_id="CAD28590.1"
                    /db_xref="GI:19578161"
                    /translation="MIVFLVFKHLFSLRLITMFFLLHFIVLINVKDFALTQGSMITPS
                    CQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGWATIFGTVHGNANSVAL
                    TQECFLKQYPQCCDCKETELECVNGDLKSVPMISNNVTLLSLKKNKIHSLPDKVFIKY
                    TKLKKIFLQHNCIRHISRKAFFGLCNLQILILDDNPITRISQRLFTGLNSLFFLSMVN
                    NYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNSTFLSCDSLTVLDLSSNTITELSPHL
                    FKDLKLLQKLDLERIEIPNINTRMFQPMKNLSHIYFKNFRYCSYAPHVRICMPLTDGI
                    SSFEDLLANNILRIFVWVIAFITCFGNLFVIGMRSFIKAENTTHAMSIKILCCADCLM
                    GVYLFFVGIFDIKYRGQYQKYALLWMESVQCRLMGFLAMLSTEVSVLLLTYLTLEKFL
                    VIVFPFSNIRPGKRQTSVILICIWMAGFLIAVIPFWNKDYFGNFYGKNGVCFPLYYDQ
                    TEDIGSKGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVAN
                    RFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTN
                    FFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLKLGVLNKITLGDSIMKPVS
                    "
     sig_peptide    1..108
BASE COUNT      579 a    415 c    395 g    660 t
ORIGIN

  Query Match             36.2%;  Score 171.8;  DB 6;  Length 2049;
  Best Local Similarity   67.4%;  Pred. No. 1.1e-29;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
              |||||  | | |||||| || ||| | || |||      || ||| ||||||| || ||||||
Db       1549 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 1608

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
              |||    |  ||||||  | |        ||||  |||  ||| |          | || ||| || |||||
Db       1609 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 1668

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTACTGATGCA 189
                  ||  || |||| ||      |||| || ||||| || ||||||||  ||  ||||||
Db       1669 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTATAGTGTTCTCTGATGCC 1728

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
                | |||||||| ||   | |||||||  |||    ||||| ||    | | ||| ||||||||
Db       1729 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 1788

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
              |  || || || || ||| ||||||||  ||       | |  ||||||||||||||| |||
Db       1789 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 1848

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
              || |||||||||||| || || |       ||||| || | | |      | | | |  |||  ||
Db       1849 ATCCTCTATACTCTCACAACCAACTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 1907

RESULT 14
AX451566
LOCUS       AX451566                2142 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 5 from Patent WO0226824.
ACCESSION   AX451566
VERSION     AX451566.1  GI:21698550
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Feder,J.N., Mintier,G., Ramanathan,C.S. and Hawken,D.R.
  TITLE     A novel human g-protein coupled receptor, hgprbmy5, expressed
            highly in brain and ovarian tissues
  JOURNAL   Patent: WO 0226824-A 5 04-APR-2002;
            Bristol-Myers Squibb Company (US)
FEATURES             Location/Qualifiers
     source          1..2142
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      613 a    435 c    408 g    686 t
ORIGIN

  Query Match             36.2%;  Score 171.8;  DB 6;  Length 2142;
  Best Local Similarity   67.4%;  Pred. No. 1.1e-29;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
              |||||  | | |||||| || ||| | || |||      || ||| ||||||| || ||||||
Db       1642 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 1701

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
              |||    |  ||||||  | |        ||||  |||  ||| |          | || ||| || |||||
Db       1702 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 1761

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTACTGATGCA 189
                  ||  || |||| ||      |||| || ||||| || ||||||||  ||  ||||||
Db       1762 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTATAGTGTTCTCTGATGCC 1821

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
                | |||||||| ||   | |||||||  |||    ||||| ||    | | ||| ||||||||
Db       1822 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 1881

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
              |  || || || || ||| ||||||||  ||       | |  ||||||||||||||| |||
Db       1882 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 1941

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
              || |||||||||||| || || |       ||||| || | | |      | | | |  |||  ||
Db       1942 ATCCTCTATACTCTCACAACCAACTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 2000

RESULT 15
AX709190
LOCUS       AX709190                2181 bp    DNA     linear   PAT 04-APR-2003
DEFINITION  Sequence 49 from Patent WO02063004.
ACCESSION   AX709190
VERSION     AX709190.1  GI:29564784
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Baughn,M.R., Tribouley,C.M., Nguyen,D.B., Thornton,M., Yao,M.G.,
            Kallick,D.A., Gandhi,A.R., Walia,N.K., Arvizu,C., Elliott,V.S.,
            Hafalia,A.J., Ramkumar,J., Pei,J., Tang,Y.T., Yue,H., Reddy,R.,
            Burford,N., Lu,D.A., Graul,R.C., Khan,F.A., Walsh,R.T., Ison,C.H.,
            Richardson,T.W., Griffin,J.A., Warren,B.A., Yang,J., Lee,E.A. and
            Harland,L.
  TITLE     G-protein coupled receptors
  JOURNAL   Patent: WO 02063004-A 49 15-AUG-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..2181
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7485090CB1"
BASE COUNT      635 a    451 c    419 g    676 t
ORIGIN

  Query Match             36.2%;  Score 171.8;  DB 6;  Length 2181;
  Best Local Similarity   67.4%;  Pred. No. 1.1e-29;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
              |||||  | | |||||| || ||| | || |||      || ||| ||||||| || ||||||
Db       1681 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 1740
```

```
Qy          70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
               ||| | ||||| | |     |||| ||| ||| |      | || ||| || |||||
Db        1741 TATATTACTATGTTCTGTTCCATTCAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 1800

Qy         130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTACTGATGCA 189
               ||  || ||| ||     |||| || ||||| || |||||||| ||  |||||||
Db        1801 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 1860

Qy         190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
               | |||||||| ||  | |||||||  |||   |||||  ||  | | ||| ||||||||
Db        1861 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 1920

Qy         250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
               |  || || ||  || ||| ||||||||  ||    | |  |||||||||||||| |||
Db        1921 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 1980

Qy         310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
               || |||||||||||| || || |     ||||| || | | |  | | | |   ||| ||
Db        1981 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 2039
```

Search completed: September 25, 2003, 13:58:08
Job time : 2168 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 12:09:23 ; Search time 222 Seconds
(without alignments)
5763.665 Million cell updates/sec

Title: US-10-049-568-1
Perfect score: 474
Sequence: 1 gcccagatttattcagtggc.....ggagatgccacctgagttaa 474

Scoring table: IDENTITY_NUC
Gpop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	474	100.0	474	22	AAP28059	Human HGRL101 G-pr
2	439.2	92.7	1089	22	AAI99557	Human expressed po
c 3	439.2	92.7	1162	23	ABK43573	DNA encoding novel
4	439.2	92.7	1191	22	AAD06507	Human CON222 G pro
5	439.2	92.7	1804	21	AAK90524	Human GPCR protein
6	439.2	92.7	1804	24	ABX73054	Human G-protein co
7	439.2	92.7	2274	25	ABK42570	Human G protein-co
8	439.2	92.7	2467	20	AAE25345	Human LGR7 long fo
9	439.2	92.7	3584	20	AAE25346	Human LGR7 short f
10	438.8	92.6	530	22	AAI99584	Human expressed po
11	438.8	92.6	530	22	AAL35644	Human musculoskele
12	438.8	92.6	530	22	ABA06471	Human cDNA SEQ ID
13	438.8	92.6	530	22	AAS28950	cDNA encoding for
14	438.8	92.6	530	22	AAS29573	Human endocrine po
15	438.8	92.6	530	22	AAS30187	DNA encoding rena
16	438.8	92.6	530	22	AAS34845	cDNA encoding nove
17	438.8	92.6	530	23	ABK43875	DNA encoding novel
18	438.8	92.6	530	24	ABY83808	Human polynucleoti
19	438.8	92.6	530	24	ABT07785	Novel human nuclei
20	438.8	92.6	530	25	ABX58632	cDNA encoding nove
21	437.6	92.3	1126	25	ACA04807	cDNA encoding huma
c 22	436.6	92.1	512	24	ABX73081	Human GPCR cDNA #4
23	353.2	74.5	420	21	AAC75735	Human ORFX ORF1290
24	171.8	36.2	1065	24	ABL40194	Human G protein-co
25	171.8	36.2	1068	22	AAS07943	Human cDNA encodin
26	171.8	36.2	1473	24	ABL40192	Human G protein-co
27	171.8	36.2	1545	24	ABL40191	Human G protein-co
28	171.8	36.2	1830	24	ABL40188	Human G protein-co
29	171.8	36.2	2142	24	ABX51947	cDNA encoding huma
30	171.8	36.2	2181	24	AAP88580	Human GCREC-1 cDNA
31	171.8	36.2	2190	24	ABL40197	Human G protein-co
32	171.8	36.2	2214	24	ABX51944	cDNA encoding huma
33	171.8	36.2	2262	24	ABL40196	Human G protein-co
34	171.8	36.2	2339	25	AAL51335	Human LGR9 G-prote
35	168.6	35.6	1121	25	ADT33349	HOVX DNA sequence
36	156.8	33.1	321	24	AAD32027	Human novel G-prot
37	154.4	32.6	636	21	AAA44932	Human secreted exp
c 38	107.2	22.6	432	24	ABX73106	Rat variant of hum
c 39	106.8	22.5	1018	22	AAH51001	Human hGPCR57 codi
c 40	106.8	22.5	1018	24	ABS70234	DNA encoding human
41	48.8	10.3	1015	22	AAS57085	cDNA encoding Dros
42	48.8	10.3	1015	23	ABL07293	Drosophila melanog
43	44	9.3	4292	21	AAF21792	Human breast and o
44	44	9.3	5644	24	ABK83490	Human cDNA differe
c 45	40.2	8.5	134499	21	AAP22286	BAC containing rep

ALIGNMENTS

RESULT 1
AAP28059
ID AAP28059 standard; cDNA; 474 BP.
XX
AC AAP28059;
XX
DT 23-MAY-2001 (first entry)
XX
DE Human HGRL101 G-protein coupled receptor coding sequence.
XX
KW Human; HGRL101; G-protein coupled receptor; infection; pain; cancer;
KW diabetes; obesity; eating disorder; asthma; Parkinson's disease;
KW hypotension; osteoporosis; myocardial infarction; migraine; allergy;
KW psychotic disorder; neurological disorder; dyskinesia; vaccine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..474
FT /tag= a
FT /product= "HGRL101"
FT /partial
XX
PN WO200114548-A2.
XX
PD 01-MAR-2001.
XX
PF 09-AUG-2000; 2000WO-EP07723.
XX
PR 19-AUG-1999; 99EP-0116345.
XX

PA (MERE) MERCK PATENT GMBH.
XX
PI Duecker K;
XX
DR WPI; 2001-226617/23.
DR P-PSDB; AAB35407.
XX
PT Novel G-coupled protein receptor, HGRL101 useful for treating diseases
PT such as microbial infections, cancers, obesity, asthma, diabetes,
PT hypotension, osteoporosis, myocardial infarction, stroke, ulcer,
PT allergy -
XX
PS Claim 5; Page 35-36; 36pp; English.
XX
CC The present invention provides the protein and coding sequences for a
CC novel human G-protein coupled receptor, designated HGRL101. The sequences
CC are useful in the diagnosis, prevention and treatment of diseases
CC including infections, pain, cancer, diabetes, anorexia, bulimia, asthma,
CC Parkinson's disease, acute heart failure, hypotension, hypertension,
CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,
CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
CC vomiting, psychotic and neurological disorders such as anxiety,
CC schizophrenia, manic depression, depression, delirium, dementia and
CC severe mental retardation, and dyskinesias including Huntington's disease
CC and Gilles dela Tourette's syndrome. The present sequence is the HGRL101
CC partial coding sequence.
XX
SQ Sequence 474 BP; 144 A; 92 C; 88 G; 150 T; 0 other;

Query Match 100.0%; Score 474; DB 22; Length 474;
Best Local Similarity 100.0%; Pred. No. 4.9e-119;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTGATGCATTATGCTGGATACCCATTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
```

RESULT 2
AAI99557
ID AAI99557 standard; cDNA; 1089 BP.
XX
AC AAI99557;
XX
DT 04-JAN-2002 (first entry)

XX
DE Human expressed polynucleotide SEQ ID NO 20.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200155387-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01310.
YX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216547.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 25-JUL-2000; 2000US-0220953.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465573/50.
DR P-PSDB; AAM99945.
XX
PT Isolated digestive system associated polypeptide for treating,
PT preventing and/or diagnosing disorders related to the digestive system
PT including digestive system cancers and also for testing and detection
PT e.g. diagnosis -
XX
PS Claim 1; SEQ ID NO 20; 509pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (anti)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1089 BP; 386 A; 186 C; 186 G; 331 T; 0 other;

Query Match 92.7%; Score 439.2; DB 22; Length 1089;
Best Local Similarity 96.8%; Pred. No. 1.8e-109;
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Qy 1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
Db 20 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 79

Qy 61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
Db 80 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 139

Qy 121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCCTTTATAGTATTT 180
Db 140 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 199

Qy 181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
Db 200 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 259

Qy 241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
Db 260 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 319

Qy 301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
Db 320 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 379

```
Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| ||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db        380 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 437

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        438 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 491


RESULT 3
ABK43573/c
ID   ABK43573 standard; cDNA; 1162 BP.
XX
AC   ABK43573;
XX
DT   05-JUN-2002  (first entry)
XX
DE   DNA encoding novel central nervous system protein #153.
XX
KW   Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW   hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW   cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW   nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW   acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW   adenocarcinoma; reproductive system disorder; testicular feminisation;
KW   endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW   respiratory disorder; renal disorder; kidney failure; blood disorder;
KW   myocardial infarction; wound healing; cell proliferation; skin aging;
KW   food additive; food preservative; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155318-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01332.
XX
PR   31-JAN-2000; 2000US-0173065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216547.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
```

```
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-581633/65.
DR   P-PSDB; AAU87243.
XX
PT   New isolated nucleic acid encoding a protein for diagnosing,
PT   preventing, treating or ameliorating medical conditions and used as
PT   food additives or preservatives -
XX
PS   Claim 1; SEQ ID No 163; 837pp; English.
XX
CC   The invention describes an isolated nucleic acid molecule (I) encoding a
CC   novel central nervous system protein. (I) and polypeptides (III) encoded
CC   by (I), are used to treat a medical conditions and in diagnosis of a
CC   pathological condition. Disorders which are diagnosed or treated include
CC   autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC   disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC   e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC   angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC   amyliotrophic lateral sclerosis, infections caused by bacteria, viruses
CC   e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC   e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC   adenocarcinomas and irritable bowel syndrome, reproductive system
CC   disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC   and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC   leukaemia, disorders involving neovascularisation e.g. malignancies,
CC   respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC   acute kidney failure and blood related disorders e.g. myocardial
CC   infarction. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities, fat content, lipid, protein,

  Query Match          92.7%;  Score 439.2;  DB 23;  Length 1162;
  Best Local Similarity  96.8%;  Pred. No. 1.8e-109;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels     2;  Gaps     1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCGCATTTATCATCATA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1080 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCGCATTTATCATCATA 1021

Qy          61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1020 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 961

Qy         121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       960 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 901

Qy         181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
               |||||||||||||||||||||||||||||||| ||||  ||||||||||||||||||||

Db       900 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 841

Qy         241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
               |||||||||||||||||||||||||||||||| | |  |||||||||||||||||||

Db       840 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTATTCTGCCATTAACAGTGCT 781

Qy         301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       780 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 721

Qy         361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       720 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 663

Qy         421 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
               |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       662 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 609


RESULT 4
AAD06507
ID   AAD06507 standard; DNA; 1191 BP.
XX
AC   AAD06507;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human CON222 G protein-coupled receptor DNA.
XX
KW   Human; G protein-coupled receptor; GPCR; CON222 protein; schizophrenia;
KW   neuroleptic; nootropic; neuroprotective; bipolar disease; psychotropic;
KW   neurological disorder; psychiatric disease; neurosis; anxiety; neuritis;
KW   attention deficit hyperactivity disorder; neurasthenia; senile dementia;
KW   affective disorder; neuropathy; Alzheimer's disease; Parkinson's disease;
KW   depression; migraine; genetic screening; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1191
FT                   /*tag= a
FT                   /product= "Human G protein-coupled receptor protein"
XX
PN   WO200131014-A2.
XX
PD   03-MAY-2001.
XX
PF   27-OCT-2000; 2000WO-US29601.
XX
PR   27-OCT-1999;   99US-0427653.
PR   27-OCT-1999;   99US-0427859.
PR   27-OCT-1999;   99US-0428020.
PR   27-OCT-1999;   99US-0428114.
PR   28-OCT-1999;   99US-0429517.
PR   28-OCT-1999;   99US-0429555.
PR   28-OCT-1999;   99US-0429676.
PR   28-OCT-1999;   99US-0429695.
PR   03-DEC-1999;   99US-0454395.
PR   12-JAN-2000; 2000US-0481794.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
```

```
XX
PI   Vogeli G,  Wood LS,  Merchant K;
XX
DR   WPI; 2001-328653/34.
DR   P-PSDB; AAE02498.
XX
PT   Seven transmembrane receptor polypeptides and polynucleotides, useful
PT   for treating neurological or psychiatric disorders, e.g. schizophrenia,
PT   as well as for identifying compounds useful for treating schizophrenia
PT   -
XX
PS   Claim 22; Page 15-16; 215pp; English.
XX
CC   The invention relates to human G protein-coupled receptor (GPCR) and
CC   their corresponding DNA molecules. GPCR is also referred as seven
CC   transmembrane receptor. G protein-coupled receptor protein is useful for
CC   treating neurological disorder, particularly schizophrenia. GPCR protein
CC   is also useful for identifying compounds useful for treating
CC   schizophrenia. These compounds are also useful for treating other
CC   neurological and psychiatric diseases, e.g. depression, anxiety, bipolar
CC   disease, affective disorders, attention deficit hyperactivity disorder/
CC   attention deficit disorder, epilepsy, neuritis, neurasthenia, neuropathy,
CC   neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile
CC   dementia. The invention also provides genetic screening procedures that
CC   entail analysing a person's genome with respect to GPCR. The vectors are
CC   useful for the recombinant production of the GPCR's. The present DNA
CC   sequence encodes human CON222 G protein-coupled receptor (GPCR) protein.
XX
SQ   Sequence 1191 BP; 340 A; 229 C; 226 G; 396 T; 0 other;

  Query Match             92.7%;  Score 439.2;  DB 22;  Length 1191;
  Best Local Similarity   96.8%;  Pred. No. 1.8e-109;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 702

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 762

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 822

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        823 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 882

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |       |||||||||||||
Db        883 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 942

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 1002

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1003 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 1060

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1061 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 1114

RESULT 5
AAZ90524
ID   AAZ90524 standard; cDNA; 1804 BP.
XX
AC   AAZ90524;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Human GPCR protein (HGPRP) encoding cDNA (clone ID 2488822).
XX
KW   Human; G protein coupled protein receptor; HGPRP; cell proliferation;
KW   neurological; immune disorder; cytostatic; anti-arteriosclerotic;
KW   anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;
KW   immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
KW   anti-Alzheimer's; anti-Parkinsonian; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200015793-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;   99WO-US20958.
XX
PR   17-SEP-1998;   98US-0156513.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Bandman O,  Lal P, 'Tang YT,  Corley NC,  Guegler KJ,  Gorgone GA;
PI   Baughn MR;
XX
DR   WPI; 2000-271432/23.
DR   P-PSDB; AAY57286.
XX
PT   Human G protein coupled protein receptor peptides useful for the
PT   prevention, diagnosis and treatment of cell proliferative, neurological
PT   and immune disorders -
XX
PS   Claim 9; Page 69-70; 71pp; English.
XX
CC   The invention provides human G protein coupled protein receptor (HGPRP)
CC   polypeptides and polynucleotides encoding them. The polypeptides can be
CC   produced by standard recombinant methodology. The polynucleotides and
CC   polypeptides may be used in the prevention, treatment and diagnosis of
CC   diseases associated with their inappropriate expression. Diseases that
CC   can be treated are cell proliferative disorders (e.g. arteriosclerosis,
CC   atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
CC   adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
CC   disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
CC   and Parkinson's disease). The anti-hGPCR antibodies may also be used as
CC   diagnostic agents for detecting the presence of HGPRP polypeptides in
CC   samples (e.g. by enzyme linked immunosorbant assay (ELISA)). Sequences
CC   AAZ90521-526 represent cDNA fragments encoding the HGPRP polypeptides
CC   (AAY57283-288).
XX
SQ   Sequence 1804 BP; 563 A; 318 C; 327 G; 596 T; 0 other;

  Query Match             92.7%;  Score 439.2;  DB 21;  Length 1804;
  Best Local Similarity   96.8%;  Pred. No. 2.1e-109;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 851

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 911

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 971

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        972 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 1031

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |       |||||||||||||
```

```
Db       1032 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTATTCTGCCCATTAACAGTGCT 1091

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 1151

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       1152 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 1209

Qy        421 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 1263


RESULT 6
ABX73054
ID   ABX73054 standard; cDNA; 1804 BP.
XX
AC   ABX73054;
XX
DT   14-MAR-2003  (first entry)
XX
DE   Human G-protein coupled receptor cDNA #4, Incyte clone 2488822CB1.
XX
KW   Human; ss; gene; G-protein coupled receptor; GPCR; neoplastic disorder;
KW   neurological disorder; immune disorder; cytostatic; pancreatic cancer;
KW   follicular carcinoma of the thyroid; leiomyoma of the uterus; epilepsy;
KW   interstitial nephritis; cancer.
XX
OS   Homo sapiens.
XX
PN   US2002106655-A1.
XX
PD   08-AUG-2002.
XX
PF   28-JUN-2001; 2001US-0895686.
XX
PR   17-SEP-1998;  98US-0156513.
XX
PA   (BAND/) BANDMAN O.
PA   (LALP/) LAL P G.
PA   (TANG/) TANG Y T.
PA   (BAUG/) BAUGHN M R.
XX
PI   Bandman O,  Lal PG,  Tang YT,  Baughn MR;
XX
DR   WPI; 2002-697866/75.
DR   P-PSDB; ABU54858.
XX
PT   New cDNAs encoding G protein coupled receptors are useful for the
PT   diagnosis, prognosis, treatment and evaluation of therapies for
PT   neoplastic, neurological and immune disorders  -
XX
PS   Claim 2; Page 31-32; 61pp; English.
XX
CC   The invention relates to an isolated cDNA encoding G-protein
CC   coupled receptor (GPCR) appearing as ABU54855-ABU54870.
CC   Also included are fragments of the cDNAs, species variants having
CC   at least 75% identity to the cDNAs, vectors comprising the cDNAs,
CC   a host cells comprising the above vectors, producing a protein
CC   (comprising culturing the above host cell under expression conditions and
CC   recovering the protein), using a cDNA to detect expression of a
CC   nucleic acid in a sample or to screen for compounds or molecules
CC   which bind to the cDNAs, using the GPCR proteins to screen
CC   compounds or molecules for ligands, using a GPCR protein to prepare and
CC   purify antibodies, an anti-GPCR antibody and using the antibody to detect
CC   expression of a GPCR protein in a sample and is diagnostic of cancer.
CC   The invention is useful for the diagnosis, prognosis, treatment
CC   and evaluation of therapies for neoplastic, neurological and immune
CC   disorders, particularly follicular carcinoma of the thyroid, leiomyoma
CC   of the uterus, pancreatic cancer, epilepsy, interstitial nephritis and
CC   immune response as a complication of cancer. The present sequence
CC   encodes a human GPCR protein of the invention.
XX
SQ   Sequence 1804 BP; 563 A; 318 C; 327 G; 596 T; 0 other;

  Query Match              92.7%;  Score 439.2;  DB 24;  Length 1804;
  Best Local Similarity    96.8%;  Pred. No. 2.1e-109;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 851

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 911

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 971

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  ||||||||||||||||||||
Db        972 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 1031

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCCATTAACAGTGCT 300
              |||||||||||||||||||||||||||||||||| | ||  ||||||||||||||||||||
Db       1032 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTATTCTGCCCATTAACAGTGCT 1091

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 1151

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       1152 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 1209

Qy        421 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 1263


RESULT 7
ABE42570
ID   ABE42570 standard; DNA; 2274 BP.
XX
AC   ABE42570;
XX
DT   04-MAR-2003  (first entry)
XX
DE   Human G protein-coupled receptor LGR7 nucleotide SEQ ID NO:622.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
KW   Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW   osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW   graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW   psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW   mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW   hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW   ulcer; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200261087-A2.
XX
PD   08-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US50107.
XX
PR   19-DEC-2000; 2000US-257144P.
XX
```

PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Rousch CL, Brown JP;
XX
DR WPI; 2003-046718/04.
DR P-PSDB; ABP81724.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating
PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
PT cancer or autoimmune diseases -
XX
PS Disclosure; Fig 1; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising:
CC (a) any one of 1601 sequences (see ABP82013 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular
CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity
CC or avidity for a particular GPCR. (I) can be used as GPCR modulators and
CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting
CC an antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunodiagnosis and immunodiagnosis. ABP42523 to ABP42869 encode
CC GPCR proteins given in ABP81675 to ABP82013, which are used in the
CC exemplification of the present invention.
XX
SQ Sequence 2274 BP; 669 A; 460 C; 421 G; 724 T; 0 other;

Query Match 92.7%; Score 439.2; DB 25; Length 2274;
Best Local Similarity 96.8%; Pred. No. 2.3e-109;
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Qy 1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
Db 1726 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 1785

Qy 61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
Db 1786 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 1845

Qy 121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
Db 1846 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 1905

Qy 181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCCACTGCTTCAGGTA 240
Db 1906 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 1965

Qy 241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTATTCTGCCATTAACAGTGCT 300
Db 1966 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 2025

Qy 301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
Db 2026 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 2085

Qy 361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
Db 2086 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 2143

Qy 421 CATCATTCATCTGGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
Db 2144 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2197

RESULT 8
AAZ25345
ID AAZ25345 standard; cDNA; 2467 BP.
XX
AC AAZ25345;
XX
DT 20-DEC-1999 (first entry)
XX
DE Human LGR7 long form nucleotide sequence.
XX
KW Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
KW extracellular leucine rich repeat region; mapping; identification; ss.
XX
OS Homo sapiens.
XX
PN WO9948921-A1.
XX
PD 30-SEP-1999.
XX
PF 25-MAR-1999; 99WO-US06573.
XX
PR 26-MAR-1998; 98US-0079501.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PA (ORGA) ORGANON NV.
XX
PI Hsueh AJW, Hsu SY, Liang S, Van Der Spek PJ;
XX
DR WPI; 1999-591074/50.
DR P-PSDB; AAY42170.
XX
PT New G-protein coupled receptors, useful for identifying their own
PT ligands -
XX
PS Claim 4; Fig 3; 54pp; English.
XX
CC The present sequence encodes the human G-protein coupled receptor
CC having extracellular leucine rich repeat regions, designated LGR7 long
CC form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
CC the receptor. The polypeptides and/or polynucleotides are also useful
CC for homologous or related genes, producing compositions that modulate
CC the expression or function of the receptors, gene therapy, mapping
CC functional regions of the receptors, studying associated physiological
CC pathways, in vivo prophylactic and therapeutic purposes, as immunogens
CC for producing antibodies, and for identifying biologically active
CC agents. The polypeptides contain a G-protein coupled seven
CC transmembrane region and a leucine rich repeat extracellular domain.
CC These regions capture and facilitate oral orientation of its ligand.
CC The proteins are also expressed in diverse tissues.
XX
SQ Sequence 2467 BP; 747 A; 487 C; 474 G; 759 T; 0 other;

Query Match 92.7%; Score 439.2; DB 20; Length 2467;
Best Local Similarity 96.8%; Pred. No. 2.3e-109;
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Qy 1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
Db 1919 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 1978

Qy 61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
Db 1979 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 2038

Qy 121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
Db 2039 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 2098

```
Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db       2099 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 2158

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |   |    |||||||||||||
Db       2159 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 2218

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 2278

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db       2279 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 2336

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       2337 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2390


RESULT 9
AAZ25346
ID   AAZ25346 standard; cDNA; 3584 BP.
XX
AC   AAZ25346;
XX
DT   20-DEC-1999  (first entry)
XX
DE   Human LGR7 short form nucleotide sequence.
XX
KW   Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
KW   extracellular leucine rich repeat region; mapping; identification; ss.
XX
OS   Homo sapiens.
XX
PN   WO9948921-A1.
XX
PD   30-SEP-1999.
XX
PF   25-MAR-1999;   99WO-US06573.
XX
PR   26-MAR-1998;   98US-0079501.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PA   (ORGA ) ORGANON NV.
XX
PI   Hsueh AJW,  Hsu SY,  Liang S,  Van Der Spek PJ;
XX
DR   WPI; 1999-591074/50.
DR   P-PSDB; AAY42171.
XX
PT   New G-protein coupled receptors, useful for identifying their own
PT   ligands -
XX
PS   Claim 4; Fig 4; 54pp; English.
XX
CC   The present sequence encodes the human G-protein coupled receptor
CC   having extracellular leucine rich repeat regions, designated LGR7 short
CC   form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
CC   the receptor. The polypeptides and/or polynucleotides are also useful
CC   for homologous or related genes, producing compositions that modulate
CC   the expression or function of the receptors, gene therapy, mapping
CC   functional regions of the receptors, studying associated physiological
CC   pathways, in vivo prophylactic and therapeutic purposes, as immunogens
CC   for producing antibodies, and for identifying biologically active
CC   agents. The polypeptides contain a G-protein coupled seven
CC   transmembrane region and a leucine rich repeat extracellular domain.
CC   These regions capture and facilitate optimal orientation of its ligand.
CC   The proteins are also expressed in diverse tissues.
XX
SQ   Sequence 3584 BP; 1124 A; 670 C; 647 G; 1139 T; 4 other;

  Query Match             92.7%;  Score 439.2;  DB 20;  Length 3584;
  Best Local Similarity   96.8%;  Pred. No. 2.6e-109;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1744 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 1803

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1804 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 1863

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1864 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 1923

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db       1924 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 1983

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |   |    |||||||||||||
Db       1984 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 2043

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2044 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 2103

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db       2104 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 2161

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       2162 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2215


RESULT 10
AAI99584
ID   AAI99584 standard; cDNA; 530 BP.
XX
AC   AAI99584;
XX
DT   04-JAN-2002  (first entry)
XX
DE   Human expressed polynucleotide SEQ ID NO 47.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200155387-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01310.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
```

PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226275.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239957.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465573/50.
DR P-PSDB; AAM99972.
XX
PT Isolated digestive system associated polypeptide for treating,
PT preventing and/ or prognosing disorders related to the digestive system
PT including digestive system cancers and also for testing and detection
PT e.g. diagnosis -
XX
PS Claim 1; SEQ ID NO 47; 509pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are

```
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC   in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC   (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

  Query Match          92.6%;  Score 438.8;  DB 22;  Length 530;
  Best Local Similarity  96.6%;  Pred. No. 1.9e-109;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              |||||||||||||||||||||||||||||||||| | |  |||||||||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTATTCTGCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473


RESULT 11
AAL35644
ID   AAL35644 standard; cDNA; 530 BP.
XX
AC   AAL35644;
XX
DT   08-JAN-2002  (first entry)
XX
DE   Human musculoskeletal system related polynucleotide SEQ ID NO 986.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein;
KW   musculoskeletal system; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155367-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01338.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
```

```
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0245474.
PR   08-NOV-2000; 2000US-0245475.
PR   08-NOV-2000; 2000US-0245476.
PR   08-NOV-2000; 2000US-0245477.
PR   08-NOV-2000; 2000US-0245478.
PR   08-NOV-2000; 2000US-0245523.
PR   08-NOV-2000; 2000US-0245524.
PR   08-NOV-2000; 2000US-0245525.
PR   08-NOV-2000; 2000US-0245526.
PR   08-NOV-2000; 2000US-0245527.
PR   08-NOV-2000; 2000US-0245528.
PR   08-NOV-2000; 2000US-0245532.
PR   08-NOV-2000; 2000US-0245603.
PR   08-NOV-2000; 2000US-0245610.
PR   08-NOV-2000; 2000US-0245611.
PR   08-NOV-2000; 2000US-0245613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX

PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-451937/48.
DR   P-PSDB; ABB04062.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders related to the musculoskeletal system including
PT   musculoskeletal cancers and also for testing and detection e.g.
PT   diagnosis -
XX
PS   Claim 1; SEQ ID NO 986; 781pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC   (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC   for preventing, treating or ameliorating medical conditions e.g. by
CC   protein or gene therapy. The genes are isolated from a range of human
CC   tissues disclosed in the specification. The nucleic acids, proteins,
CC   antibodies and (ant)agonists are useful in the diagnosis, treatment
CC   and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC   other cancers of the adrenal gland, bone, bone marrow, breast,
CC   gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC   healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC   and (f) infectious diseases such as viral, bacterial, fungal and
CC   parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

  Query Match          92.6%;  Score 438.8;  DB 22;  Length 530;
  Best Local Similarity  96.6%;  Pred. No. 1.9e-109;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||| ||||  ||||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              |||||||||||||||||||||||||||||||| |  |  ||||||||||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTATTCTGCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||:||  |||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAKAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473


RESULT 12
ABA06471
```

ID ABA06471 standard; cDNA; 530 BP.
XX
AC ABA06471;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human cDNA SEQ ID NO: 137.
XX
KW Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200154474-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01349.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 24-FEB-2000; 2000US-184664P.
PR 02-MAR-2000; 2000US-186350P.
PR 16-MAR-2000; 2000US-189874P.
PR 17-MAR-2000; 2000US-190076P.
PR 18-APR-2000; 2000US-198123P.
PR 19-MAY-2000; 2000US-205515P.
PR 07-JUN-2000; 2000US-209467P.
PR 28-JUN-2000; 2000US-214886P.
PR 30-JUN-2000; 2000US-215135P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225213P.
PR 14-AUG-2000; 2000US-225214P.
PR 14-AUG-2000; 2000US-225266P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 14-AUG-2000; 2000US-225759P.
PR 18-AUG-2000; 2000US-226279P.
PR 22-AUG-2000; 2000US-226681P.
PR 22-AUG-2000; 2000US-226868P.
PR 22-AUG-2000; 2000US-227182P.
PR 23-AUG-2000; 2000US-227009P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 06-SEP-2000; 2000US-230437P.
PR 06-SEP-2000; 2000US-230438P.
PR 08-SEP-2000; 2000US-231242P.
PR 08-SEP-2000; 2000US-231243P.
PR 08-SEP-2000; 2000US-231244P.
PR 08-SEP-2000; 2000US-231413P.
PR 08-SEP-2000; 2000US-231414P.
PR 08-SEP-2000; 2000US-232080P.
PR 08-SEP-2000; 2000US-232081P.
PR 12-SEP-2000; 2000US-231968P.
PR 14-SEP-2000; 2000US-232397P.
PR 14-SEP-2000; 2000US-232398P.
PR 14-SEP-2000; 2000US-232399P.
PR 14-SEP-2000; 2000US-232400P.
PR 14-SEP-2000; 2000US-232401P.
PR 14-SEP-2000; 2000US-233063P.
PR 14-SEP-2000; 2000US-233064P.
PR 14-SEP-2000; 2000US-233065P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234977P.
PR 25-SEP-2000; 2000US-234998P.
PR 26-SEP-2000; 2000US-235484P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-235836P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 13-OCT-2000; 2000US-239937P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241221P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241786P.
PR 20-OCT-2000; 2000US-241787P.
PR 20-OCT-2000; 2000US-241808P.
PR 20-OCT-2000; 2000US-241809P.
PR 20-OCT-2000; 2000US-241826P.
PR 01-NOV-2000; 2000US-244617P.
PR 08-NOV-2000; 2000US-246474P.
PR 08-NOV-2000; 2000US-246475P.
PR 08-NOV-2000; 2000US-246476P.
PR 08-NOV-2000; 2000US-246477P.
PR 08-NOV-2000; 2000US-246478P.
PR 08-NOV-2000; 2000US-246523P.
PR 08-NOV-2000; 2000US-246524P.
PR 08-NOV-2000; 2000US-246525P.
PR 08-NOV-2000; 2000US-246526P.
PR 08-NOV-2000; 2000US-246527P.
PR 08-NOV-2000; 2000US-246528P.
PR 08-NOV-2000; 2000US-246532P.
PR 08-NOV-2000; 2000US-246609P.
PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246611P.
PR 08-NOV-2000; 2000US-246613P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.
PR 17-NOV-2000; 2000US-249209P.
PR 17-NOV-2000; 2000US-249210P.
PR 17-NOV-2000; 2000US-249211P.
PR 17-NOV-2000; 2000US-249212P.
PR 17-NOV-2000; 2000US-249213P.
PR 17-NOV-2000; 2000US-249214P.
PR 17-NOV-2000; 2000US-249215P.
PR 17-NOV-2000; 2000US-249216P.
PR 17-NOV-2000; 2000US-249217P.
PR 17-NOV-2000; 2000US-249218P.
PR 17-NOV-2000; 2000US-249244P.
PR 17-NOV-2000; 2000US-249245P.
PR 17-NOV-2000; 2000US-249264P.
PR 17-NOV-2000; 2000US-249265P.
PR 17-NOV-2000; 2000US-249297P.
PR 17-NOV-2000; 2000US-249299P.
PR 17-NOV-2000; 2000US-249300P.
PR 01-DEC-2000; 2000US-250160P.
PR 01-DEC-2000; 2000US-250391P.

```
PR   05-DEC-2000; 2000US-251030P.
PR   05-DEC-2000; 2000US-251988P.
PR   05-DEC-2000; 2000US-256719P.
PR   06-DEC-2000; 2000US-251479P.
PR   08-DEC-2000; 2000US-251856P.
PR   08-DEC-2000; 2000US-251868P.
PR   08-DEC-2000; 2000US-251869P.
PR   08-DEC-2000; 2000US-251989P.
PR   08-DEC-2000; 2000US-251990P.
PR   11-DEC-2000; 2000US-254097P.
PR   05-JAN-2001; 2001US-259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-476161/51.
DR   P-PSDB; ABB10249.
XX
PT   Isolated nucleic acid molecule encoding an inflammation-associated
PT   polypeptide is used in preventing, treating or ameliorating a medical
PT   condition -
XX
PS   Claim 1; SEQ ID NO: 137; 859pp + Sequence Listing; English;
XX
CC   The present invention provides human cDNAs, proteins and related genomic
CC   DNAs. These can be used in the treatment of neural, immune system,
CC   muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC   renal and proliferative disorders and inflammation. The present sequence
CC   is a cDNA of the invention.
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

  Query Match             92.6%;  Score 438.8;  DB 22;  Length 530;
  Best Local Similarity   96.6%;  Pred. No. 1.9e-109;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTACGCAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              |||||||||||||||||||||||||||||||| |  |  ||||||||||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||:|||  ||||||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473


RESULT 13
AAS28950
ID   AAS28950 standard; cDNA; 530 BP.
XX
AC   AAS28950;
XX
DT   21-NOV-2001  (first entry)
XX
DE   cDNA encoding for human uterine motility-association polypeptide #15.
XX
KW   Human; uterine motility-association disorder; uterus; pregnancy;
KW   labour; menstrual cycle; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155201-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01317.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   15-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
```

```
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-488777/53.
DR   P-PSDB; AAU18108.
XX
PT   Isolated polypeptide and nucleic acid molecules for treating,
PT   preventing and/or prognosing disorders related to uterine motility
PT   e.g. disorders associated with pregnancy and the menstrual cycle -
XX
PS   Claim 4; SEQ ID NO 25; 524pp; English.
XX
CC   The present invention relates to the isolation of novel human
CC   uterine motility-association polypeptides (AAU18094-AAU18152),
CC   and cDNA and genomic sequences encoding for these polypeptides.
CC   The sequences of the invention are useful in the diagnosis,
CC   treatment, prevention and/or prognosis of diseases associated
CC   with uterine motility such as pregnancy and labour, and menstrual
CC   disorders. The polynucleotide sequences of the invention are also
CC   useful in gene therapy. AAS28936-AAS28994 represent cDNA sequences
CC   encoding for novel human uterine motility-association polypeptides.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

  Query Match             92.6%;  Score 438.8;  DB 22;  Length 530;
  Best Local Similarity   96.6%;  Pred. No. 1.9e-109;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||   |   |       |||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||:|||  |||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473
```

RESULT 14
AAS29573
ID AAS29573 standard; cDNA; 530 BP.
XX
AC AAS29573;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human endocrine polypeptide encoding cDNA SEQ ID No 73.
XX
KW Endocrine protein; human; mouse; rabbit; goat; horse; food additive;
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; as;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility.
XX
OS Homo sapiens.
XX
PN WO200155364-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01308.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.

```
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-451936/48.
DR   P-PSDB; AAU18344.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders of the endocrine system such as reproductive disorders,
PT   endocrine cancers and also for testing and detection e.g. diagnosis -
XX
PS   Claim 1; SEQ ID No 73; 604pp; English.
XX
CC   Sequences AAS29511-AAS29736 represent cDNA molecules, which encode the
CC   endocrine polypeptides of the invention. Endocrine polypeptides and their
CC   associated polynucleotides of the invention are useful in the diagnosis,
CC   treatment and prevention of various types of disorders in e.g. humans,
CC   mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC   pathological condition can be determined by determining the presence or
CC   absence of a mutation in an endocrine polynucleotide. The treatable
CC   disorders include autoimmune diseases such as rheumatoid arthritis,
CC   hyperproliferative disorders such as neoplasms of the breast or liver,
CC   cardiovascular disorders such as cardiac arrest, cerebrovascular
CC   disorders such as cerebral ischaemia, nervous system disorders such as
CC   Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC   ocular disorders such as corneal infection, endocrine disorders such as
CC   premature labour and infertility, gastrointestinal disorders such as
CC   Crohn's disease, renal disorders such as glomerulonephritis and
CC   respiratory disorders such as asthma. The polypeptides can also be used
CC   to aid wound healing, to prevent skin aging due to sunburn, to maintain
CC   organs before transplantation, to regenerate tissues and in chemotaxis.
CC   The polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO

  Query Match             92.6%;  Score 438.8;  DB 22;  Length 530;
  Best Local Similarity   96.6%;  Pred. No. 1.9e-109;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |     |||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||:|||  |||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473


RESULT 15
AAS30187
ID   AAS30187 standard; cDNA; 530 BP.
XX
AC   AAS30187;
XX
DT   21-NOV-2001  (first entry)
XX
DE   DNA encoding  renal and cardiovascular-associated protein, Seq ID 33.
XX
KW   Human; antiinflammatory; neuroprotective; immunomodulator; vulnerary;
KW   cardiovascular; cytostatic; nephrotropic; antianaemic; nephritis;
KW   immunosuppressive; kidney disorder; renal failure; hypertension;
KW   cardiovascular disorder; myocardial infarction; blood disorder; anaemia;
KW   blood coagulation disorder; electrolyte imbalance disorder; cancer;
KW   hyponatraemia; hyperkalaemia; neoplastic disorder; nephtoma;
KW   autoimmune disease; inflammatory disease; reproductive system disorder;
KW   endocrine disorder; neural activity; neurological disorder;
KW   wound healing; respiratory disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155328-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01359.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
```

PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226651.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488787/53.
DR P-PSDB; AAU18666.
XX
PT New polynucleotides and polypeptides, useful for diagnosing, treating,
PT preventing or prognosing e.g. kidney, cardiovascular, blood,
PT electrolyte imbalance or neoplastic disorders, autoimmune diseases,
PT cancers -
XX
PS Claim 1; SEQ ID No 33; 506pp; English.
XX
CC The invention relates to novel nucleic acids and polypeptides useful for
CC diagnosing, treating, preventing and/or prognosing disorders related to
CC these polypeptides. The polynucliectides are especially useful in the
CC diagnosis, prognosis, prevention and/or treatment of diseases which
CC include kidney disorders (e.g. renal failure or nephritis),
CC cardiovascular disorders (e.g. hypertension or myocardial infarction),
CC blood disorders (e.g. anaemia or blood coagulation disorders),
CC electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),
CC neoplastic disorders (e.g. nephtoma or renal cell cancer), autoimmune
CC diseases, cancers, inflammatory diseases, reproductive system
CC disorders, endocrine disorders, neural activity and neurological
CC disorders, wound healing and respiratory disorders. AA330165-AA330251
CC represent the novel human renal and cardiovascular-associated nucleic
CC acid sequences of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at:
CC ftp.wipo.int/pub/published_pct_sequences.
XX

```
SQ   Sequence 530 BP; 158 A; 105 C; 95 G; 171 T; 1 other;

  Query Match           92.6%;  Score 438.8;  DB 22;  Length 530;
  Best Local Similarity  96.6%;  Pred. No. 1.9e-109;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTATTCTGCCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||:|||  |||||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473


Search completed: September 25, 2003, 13:21:51
Job time : 225 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 13:58:14 ; Search time 190 Seconds
(without alignments)
6210.625 Million cell updates/sec

Title: US-10-049-568-1
Perfect score: 474
Sequence: 1 gccagatttattcagtggc.....ggagatgccacctgagttaa 474

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1678620 seqs, 1244745471 residues

Total number of hits satisfying chosen parameters: 3357240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	439.2	92.7	1089	14	US-10-073-885-20	Sequence 20, Appl
2	439.2	92.7	1804	10	US-09-895-686-10	Sequence 10, Appl
3	439.2	92.7	2274	14	US-10-225-567A-622	Sequence 622, App
4	438.8	92.6	530	9	US-09-764-853-137	Sequence 137, App
5	438.8	92.6	530	10	US-09-764-877-986	Sequence 986, App
6	438.8	92.6	530	11	US-09-989-442-33	Sequence 33, Appl
7	438.8	92.6	530	12	US-09-764-886-18	Sequence 18, Appl
8	438.8	92.6	530	14	US-10-073-865-25	Sequence 25, Appl
9	438.8	92.6	530	14	US-10-103-313-89	Sequence 89, Appl
10	438.8	92.6	530	14	US-10-073-885-47	Sequence 47, Appl
11	437.6	92.3	1126	14	US-10-313-542-255	Sequence 255, App
c 12	436.6	92.1	612	10	US-09-895-686-37	Sequence 37, Appl
13	171.8	36.2	1068	12	US-10-321-807-17	Sequence 17, Appl
14	171.8	36.2	2049	10	US-09-928-175-11	Sequence 11, Appl
15	171.8	36.2	2212	11	US-09-965-536A-5	Sequence 5, Appli
16	171.8	36.2	2193	10	US-09-928-175-6	Sequence 6, Appli
17	171.8	36.2	2214	11	US-09-965-536A-1	Sequence 1, Appli
18	171.8	36.2	2265	10	US-09-928-175-1	Sequence 1, Appli
19	171.8	36.2	2436	14	US-10-229-735-2	Sequence 2, Appli
20	171.8	36.2	2838	14	US-10-222-668-1	Sequence 1, Appli
21	156.8	33.1	321	11	US-09-930-312-1	Sequence 1, Appli
22	146.2	30.8	2214	10	US-09-928-175-19	Sequence 19, Appl
23	146.2	30.8	2539	14	US-10-229-735-4	Sequence 4, Appli
24	130.4	27.5	719	12	US-09-814-353-18321	Sequence 18321, A
25	113	23.8	677	12	US-09-814-353-5651	Sequence 5651, Ap
26	113	23.8	677	12	US-09-814-353-11937	Sequence 11937, A
c 27	107.2	22.6	432	10	US-09-895-686-62	Sequence 62, Appl
c 28	106.8	22.5	1018	11	US-09-782-974C-65	Sequence 65, Appl
29	48.8	10.3	1015	14	US-10-270-333-20	Sequence 20, Appl
30	44	9.3	4292	14	US-10-102-806-179	Sequence 179, App
31	44	9.3	5516	14	US-10-153-668-271	Sequence 271, App
32	39.8	8.4	113515	12	US-10-311-455-2147	Sequence 2147, Ap
33	38.8	8.2	12138	12	US-10-311-455-1602	Sequence 1602, Ap
c 34	38	8.0	444	13	US-10-027-632-62629	Sequence 62629, A
c 35	38	8.0	444	13	US-10-027-632-62630	Sequence 62630, A
c 36	38	8.0	444	13	US-10-027-632-306540	Sequence 306540,
c 37	38	8.0	444	13	US-10-027-632-306541	Sequence 306541,
c 38	37.8	8.0	1265	9	US-09-810-936-169	Sequence 169, App
c 39	37.8	8.0	1265	10	US-09-429-755-169	Sequence 169, App
c 40	37.8	8.0	1265	10	US-09-924-400-169	Sequence 169, App
c 41	37.8	8.0	1265	14	US-10-212-679-169	Sequence 169, App
c 42	37.8	8.0	1646	9	US-09-810-936-290	Sequence 290, App
c 43	37.8	8.0	1646	10	US-09-429-755-290	Sequence 290, App
c 44	37.8	8.0	1646	10	US-09-924-400-290	Sequence 290, App
c 45	37.8	8.0	1646	14	US-10-212-679-290	Sequence 290, App

ALIGNMENTS

RESULT 1
US-10-073-885-20
; Sequence 20, Application US/10073885
; Publication No. US20030096346A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ203C1
; CURRENT APPLICATION NUMBER: US/10/073,885
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-073-885-20

Query Match 92.7%; Score 439.2; DB 14; Length 1089;
Best Local Similarity 96.8%; Pred. No. 7.8e-111;
Matches 459; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

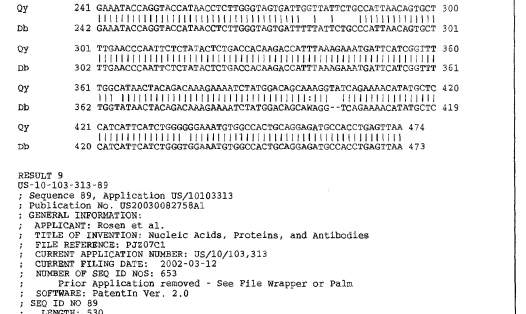
```
Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 79

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 139

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 199

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              ||||||||||||||||||||||||||||||||||||| |||| ||||||||||||||||
Db        200 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 259
```

```
Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||    |   |      ||||||||||||
Db        260 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 319

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 379

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        380 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 437

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        438 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 491


RESULT 2
US-09-895-686-10
; Sequence 10, Application US/09895686
; Patent No. US20020106655A1
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT:  Lal, Preeti
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION: HUMAN GPCR PROTEINS
;  FILE REFERENCE: PC-0044 CIP
;  CURRENT APPLICATION NUMBER: US/09/895,686
;  CURRENT FILING DATE:  2001-06-28
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PERL Program
; SEQ ID NO 10
;   LENGTH: 1804
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020106655A1 2488822CB1
US-09-895-686-10

  Query Match             92.7%;  Score 439.2;  DB 10;  Length 1804;
  Best Local Similarity   96.8%;  Pred. No. 1e-110;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 851

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 911

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 971

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||   ||||||||||||||||||
Db        972 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 1031

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||    |   |      ||||||||||||
Db       1032 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 1091

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 1151

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1152 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 1209

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1210 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 1263


RESULT 3
US-10-225-567A-622
; Sequence 622, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPT
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 622
;   LENGTH: 2274
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-622

  Query Match             92.7%;  Score 439.2;  DB 14;  Length 2274;
  Best Local Similarity   96.8%;  Pred. No. 1.2e-110;
  Matches  459;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 1785

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 1845

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 1905

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||   ||||||||||||||||||
Db       1906 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 1965

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||    |   |      ||||||||||||
Db       1966 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 2025

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 2085

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db       2086 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 2143

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       2144 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 2197


RESULT 4
US-09-764-853-137
; Sequence 137, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJ206
```

```
;      MEDIUM TYPE:  Diskette 3.50", 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Processor (WordPerfect 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/853,659A
;      FILING DATE:  Unknown
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  none
;      FILING DATE:  n/a
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8967 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  double stranded
;      TOPOLOGY:  linear
US-08-853-659A-9

  Query Match              8.1%;  Score 38.4;  DB 2;  Length 8967;
  Best Local Similarity   62.5%;  Pred. No. 0.53;
  Matches   60;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy         98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
              ||| ||||||||| |   |  ||||| |  | ||  |||  |||||     || || | |  |
Db       1844 AAATTGCCATAAAAAGAGATGAAAAAAAGTATTTAGTATAAAAAATCTTGTTCATCGATA 1903

Qy        158 AACGTTTTTTCTTTATAGTATTTACTGATGCATTAT 193
              ||    |||||  | ||| | |||||| |   | |||
Db       1904 AAGAGTTTTTTATGATATTTTTTACTCTTATACTAT 1939


RESULT 5
US-08-853-659A-64/c
; Sequence 64, Application US/08853659A
; Patent No. 5925522
;  GENERAL INFORMATION:
;    APPLICANT:  Wong, K.K.; Saffer, J.D.
;    TITLE OF INVENTION:  A Salmonella Sequence, Methods Of Detection
;    TITLE OF INVENTION:  Of A
;    TITLE OF INVENTION:  Salmonella Sequence, And Methods Of Detection Of Salmonella
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Paul W. Zimmerman
;      ADDRESSEE:  Intellectual Property Services
;      ADDRESSEE:  Battelle Memorial Institute
;      ADDRESSEE:  PNNL P.O. Box 999
;      STREET:  Washington Way
;      CITY:  Richland
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  99352
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette 3.50", 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Processor (WordPerfect 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/853,659A
;      FILING DATE:  Unknown
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  none
;      FILING DATE:  n/a
;  INFORMATION FOR SEQ ID NO:  64:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8967 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
US-08-853-659A-64

  Query Match              8.1%;  Score 38.4;  DB 2;  Length 8967;
  Best Local Similarity   62.5%;  Pred. No. 0.53;
  Matches   60;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy         98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
              ||| |||||||| |   |  ||||| |  | ||  |||  |||||     || || | |  |
Db       7124 AAATTGCCATAAAAAGAGATGAAAAAAAGTATTTAGTATAAAAAATCTTGTTCATCGATA 7065

Qy        158 AACGTTTTTTCTTTATAGTATTTACTGATGCATTAT 193
              ||    |||||  | ||| | |||||| |   | |||
Db       7064 AAGAGTTTTTTATGATATTTTTTACTCTTATACTAT 7029


RESULT 6
US-08-853-659A-67
; Sequence 67, Application US/08853659A
; Patent No. 5925522
;  GENERAL INFORMATION:
;    APPLICANT:  Wong, K.K.; Saffer, J.D.
;    TITLE OF INVENTION:  A Salmonella Sequence, Methods Of Detection
;    TITLE OF INVENTION:  Of A
;    TITLE OF INVENTION:  Salmonella Sequence, And Methods Of Detection Of Salmonell
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Paul W. Zimmerman
;      ADDRESSEE:  Intellectual Property Services
;      ADDRESSEE:  Battelle Memorial Institute
;      ADDRESSEE:  PNNL P.O. Box 999
;      STREET:  Washington Way
;      CITY:  Richland
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  99352
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette 3.50", 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Processor (WordPerfect 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/853,659A
;      FILING DATE:  Unknown
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  none
;      FILING DATE:  n/a
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8967 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
US-08-853-659A-67

  Query Match              8.1%;  Score 38.4;  DB 2;  Length 8967;
  Best Local Similarity   38.5%;  Pred. No. 0.53;
  Matches   37;  Conservative   23;  Mismatches   36;  Indels    0;  Gaps    0;

Qy         98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
              ||| :||||:|| |   |  :|||| |  | |:  ||:  |||||     :| :| : |  |
Db       1844 AAAUUGCCAUAAAAAGAGAUGAAAAAAAGUAUUUAGUAUAAAAAAUCUUGUUCAUCGAUA 1903

Qy        158 AACGTTTTTTCTTTATAGTATTTACTGATGCATTAT 193
              ||    :::::  : |:| : :::||: :   | |:
Db       1904 AAGAGUUUUUUAUGAUAUUUUUUACUCUUAUACUAU 1939


RESULT 7
US-08-853-659A-2/c
; Sequence 2, Application US/08853659A
; Patent No. 5925522
;  GENERAL INFORMATION:
;    APPLICANT:  Wong, K.K.; Saffer, J.D.
;    TITLE OF INVENTION:  A Salmonella Sequence, Methods Of Detection
;    TITLE OF INVENTION:  Of A
```

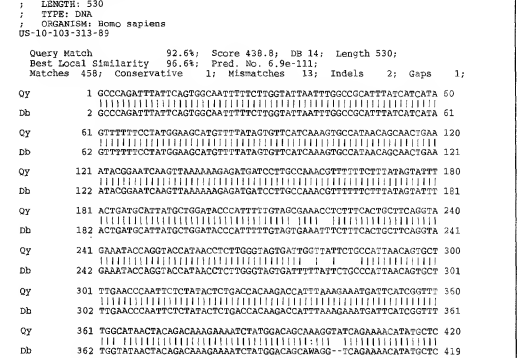
```
;    TITLE OF INVENTION:  Salmonella Sequence, And Methods Of Detection Of Salmonella
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Paul W. Zimmerman
;      ADDRESSEE:  Intellectual Property Services
;      ADDRESSEE:  Battelle Memorial Institute
;      ADDRESSEE:  PNNL P.O. Box 999
;      STREET:  Washington Way
;      CITY:  Richland
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  99352
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette 3.50", 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Processor (WordPerfect 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/853,659A
;      FILING DATE:  Unknown
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  none
;      FILING DATE:  n/a
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24701 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  double stranded
;      TOPOLOGY:  linear
US-08-853-659A-2

  Query Match             8.1%;  Score 38.4;  DB 2;  Length 24701;
  Best Local Similarity   62.5%;  Pred. No. 0.68;
  Matches   60;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy         98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
               ||| |||||||| |  |  ||||| |  | ||  |||  |||||     || || | |  |
Db      22858 AAATTGCCATAAAAAGAGATGAAAAAAAGTATTTAGTATAAAAATCTTGTTCATCGATA 22799

Qy        158 AACGTTTTTTCTTTATAGTATTTACTGATGCATTAT 193
               ||   |||||  | ||| | ||||||  |  | |||
Db      22798 AAGAGTTTTTTATGATATTTTTTACTCTTATACTAT 22763


RESULT 8
US-08-853-659A-3
; Sequence 3, Application US/08853659A
; Patent No. 5925522
;  GENERAL INFORMATION:
;    APPLICANT:  Wong, K.K.; Saffer, J.D.
;    TITLE OF INVENTION:  A Salmonella Sequence, Methods Of Detection
;    TITLE OF INVENTION:  Of A
;    TITLE OF INVENTION:  Salmonella Sequence, And Methods Of Detection Of Salmonella
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Paul W. Zimmerman
;      ADDRESSEE:  Intellectual Property Services
;      ADDRESSEE:  Battelle Memorial Institute
;      ADDRESSEE:  PNNL P.O. Box 999
;      STREET:  Washington Way
;      CITY:  Richland
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  99352
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette 3.50", 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Processor (WordPerfect 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/853,659A
;      FILING DATE:  Unknown
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  none
;      FILING DATE:  n/a
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24701 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  double stranded
;      TOPOLOGY:  linear
US-08-853-659A-3

  Query Match             8.1%;  Score 38.4;  DB 2;  Length 24701;
  Best Local Similarity   62.5%;  Pred. No. 0.68;
  Matches   60;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy         98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
               ||| |||||||| |  |  ||||| |  | ||  |||  |||||     || || | |  |
Db       1844 AAATTGCCATAAAAAGAGATGAAAAAAAGTATTTAGTATAAAAATCTTGTTCATCGATA 1903

Qy        158 AACGTTTTTTCTTTATAGTATTTACTGATGCATTAT 193
               ||   |||||  | ||| | ||||||  |  | |||
Db       1904 AAGAGTTTTTTATGATATTTTTTACTCTTATACTAT 1939


RESULT 9
US-08-853-659A-60/c
; Sequence 60, Application US/08853659A
; Patent No. 5925522
;  GENERAL INFORMATION:
;    APPLICANT:  Wong, K.K.; Saffer, J.D.
;    TITLE OF INVENTION:  A Salmonella Sequence, Methods Of Detection
;    TITLE OF INVENTION:  Of A
;    TITLE OF INVENTION:  Salmonella Sequence, And Methods Of Detection Of Salmonella
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Paul W. Zimmerman
;      ADDRESSEE:  Intellectual Property Services
;      ADDRESSEE:  Battelle Memorial Institute
;      ADDRESSEE:  PNNL P.O. Box 999
;      STREET:  Washington Way
;      CITY:  Richland
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  99352
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette 3.50", 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Processor (WordPerfect 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/853,659A
;      FILING DATE:  Unknown
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  none
;      FILING DATE:  n/a
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24701 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
US-08-853-659A-60

  Query Match             8.1%;  Score 38.4;  DB 2;  Length 24701;
  Best Local Similarity   62.5%;  Pred. No. 0.68;
  Matches   60;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy         98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
               ||| |||||||| |  |  ||||| |  | ||  |||  |||||     || || | |  |
Db      22858 AAATTGCCATAAAAAGAGATGAAAAAAAGTATTTAGTATAAAAATCTTGTTCATCGATA 22799
```

```
;  PRIOR APPLICATION NUMBER: 60/241,808
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,826
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,786
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,221
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/246,475
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/231,243
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/233,065
;  PRIOR FILING DATE: 2000-09-14

  Query Match             92.6%;  Score 438.8;  DB 11;  Length 530;
  Best Local Similarity   96.6%;  Pred. No. 6.9e-111;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |    ||||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| ||||||||||||||||||||||||||||||||:|||  ||||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473

RESULT 7
US-09-764-886-18
; Sequence 18, Application US/09764886
; Publication No. US20030139327A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PZ02
;  CURRENT APPLICATION NUMBER: US/09/764,886
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 88
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-886-18

  Query Match             92.6%;  Score 438.8;  DB 12;  Length 530;
  Best Local Similarity   96.6%;  Pred. No. 6.9e-111;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |    ||||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| ||||||||||||||||||||||||||||||||:|||  ||||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473

RESULT 8
US-10-073-865-25
; Sequence 25, Application US/10073865
; Publication No. US20030044904A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ09C1
;  CURRENT APPLICATION NUMBER: US/10/073,865
;  CURRENT FILING DATE:   2002-02-14
;  Prior Application removed - See file Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 154
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-073-865-25

  Query Match             92.6%;  Score 438.8;  DB 14;  Length 530;
  Best Local Similarity   96.6%;  Pred. No. 6.9e-111;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241
```



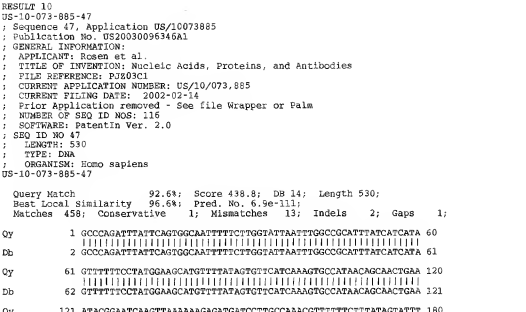
RESULT 9
US-10-103-313-89
; Sequence 89, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJS07C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-103-313-89

Query Match 92.6%; Score 438.8; DB 14; Length 530;
Best Local Similarity 96.6%; Pred. No. 6.9e-111;
Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;



RESULT 10
US-10-073-885-47
; Sequence 47, Application US/10073885
; Publication No. US20030096346A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJS03C1
; CURRENT APPLICATION NUMBER: US/10/073,885
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-073-885-47

Query Match 92.6%; Score 438.8; DB 14; Length 530;
Best Local Similarity 96.6%; Pred. No. 6.9e-111;
Matches 458; Conservative 1; Mismatches 13; Indels 2; Gaps 1;



RESULT 11
US-10-313-542-255
; Sequence 255, Application US/10313542
; Publication No. US20030120057A1
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATE
; FILE REFERENCE: PA-0013 DS
; CURRENT APPLICATION NUMBER: US/10/313,542
; CURRENT FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: US/09/495,050
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: 1999-02-01

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 13:00:53 ; Search time 55 Seconds
(without alignments)
3803.920 Million cell updates/sec

Title: US-10-049-568-1
Perfect score: 474
Sequence: 1 gcccagatttattcagtggc.....ggagatgccacctgagttaa 474

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

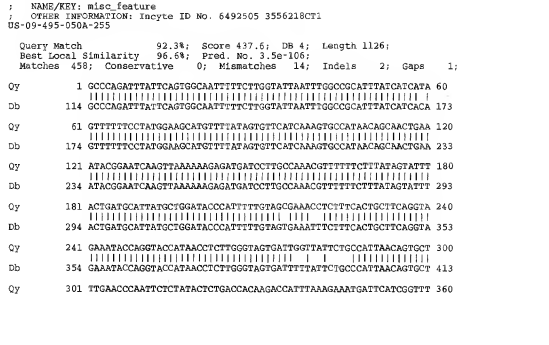
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	437.6	92.3	1126	4	US-09-495-050A-255	Sequence 255, App
c 2	38.4	8.1	1797	2	US-08-853-659A-28	Sequence 28, Appl
c 3	38.4	8.1	8967	2	US-08-853-659A-6	Sequence 6, Appli
4	38.4	8.1	8967	2	US-08-853-659A-9	Sequence 9, Appli
c 5	38.4	8.1	8967	2	US-08-853-659A-64	Sequence 64, Appl
6	38.4	8.1	8967	2	US-08-853-659A-67	Sequence 67, Appl
c 7	38.4	8.1	24701	2	US-08-853-659A-2	Sequence 2, Appli
8	38.4	8.1	24701	2	US-08-853-659A-3	Sequence 3, Appli
c 9	38.4	8.1	24701	2	US-08-853-659A-60	Sequence 60, Appl
10	38.4	8.1	24701	2	US-08-853-659A-61	Sequence 61, Appl
c 11	37.8	8.0	1265	3	US-08-991-789A-169	Sequence 169, App
c 12	37.8	8.0	1265	4	US-09-062-451-169	Sequence 169, App
c 13	37.8	8.0	1265	4	US-09-598-325-169	Sequence 169, App
c 14	37.8	8.0	1265	4	US-09-289-198-169	Sequence 169, App
c 15	37.8	8.0	1557	4	US-09-134-001C-1232	Sequence 1232, Ap
c 16	37.8	8.0	1646	3	US-08-991-789A-290	Sequence 290, App
c 17	37.8	8.0	1646	4	US-09-062-451-290	Sequence 290, App
c 18	37.8	8.0	1646	4	US-09-289-198-290	Sequence 290, App
c 19	37.2	7.8	1440	4	US-09-205-258-81	Sequence 81, Appl
20	36	7.6	2166	4	US-09-601-198-95	Sequence 95, Appl
c 21	35.4	7.5	836	3	US-09-352-990-7	Sequence 7, Appli
22	34.2	7.2	87350	3	US-08-781-891-79	Sequence 79, Appl
c 23	34.2	7.2	87350	3	US-08-781-891-79	Sequence 79, Appl
24	34.2	7.2	87350	4	US-09-618-166-79	Sequence 79, Appl
c 25	34.2	7.2	87350	4	US-09-618-166-79	Sequence 79, Appl
26	34.2	7.2	87543	4	US-09-791-211-3	Sequence 3, Appli
c 27	34.2	7.2	87543	4	US-09-791-211-3	Sequence 3, Appli
c 28	34.2	7.2	1830121	4	US-09-557-884-1	Sequence 1, Appli
c 29	34.2	7.2	1830121	4	US-09-643-990A-1	Sequence 1, Appli
30	33.6	7.1	72604	4	US-09-268-992-7	Sequence 7, Appli
31	33.6	7.1	72604	4	US-09-657-474-7	Sequence 7, Appli
32	33.2	7.0	1830121	4	US-09-557-884-1	Sequence 1, Appli
33	33.2	7.0	1830121	4	US-09-643-990A-1	Sequence 1, Appli
34	33	7.0	615	3	US-08-998-416-186	Sequence 186, App
35	33	7.0	636	3	US-08-998-416-1137	Sequence 1137, Ap
36	33	7.0	837	3	US-08-998-416-288	Sequence 288, App
c 37	33	7.0	15462	3	US-09-073-492-1	Sequence 1, Appli
c 38	32.6	6.9	498	4	US-08-936-165A-53	Sequence 53, Appl
39	32.2	6.8	8700	2	US-08-392-625-16	Sequence 16, Appl
40	32.2	6.8	8700	2	US-08-466-961A-16	Sequence 16, Appl
41	32.2	6.8	8700	2	US-08-645-193B-18	Sequence 18, Appl
42	31.8	6.7	665	2	US-08-883-795A-36	Sequence 36, Appl
43	31.8	6.7	1300	4	US-08-936-165A-253	Sequence 253, Appl
44	31.8	6.7	3001	4	US-09-539-333D-215	Sequence 215, App
45	31.8	6.7	3677	2	US-08-390-888A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-495-050A-255
; Sequence 255, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIAT
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 255
; LENGTH: 1126
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6492505 3556218CT1
US-09-495-050A-255

Query Match 92.3%; Score 437.6; DB 4; Length 1126;
Best Local Similarity 96.6%; Pred. No. 3.5e-106;
Matches 458; Conservative 0; Mismatches 14; Indels 2; Gaps 1;



```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       414 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 473

Qy       361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| ||||||||||||||||||||||||||||||||||||||||  |||||||||||||||
Db       474 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 531

Qy       421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
             ||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       532 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 585
```

RESULT 2
US-08-853-659A-28/c
; Sequence 28, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
;   APPLICANT: Wong, K.K.; Saffer, J.D.
;   TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
;   TITLE OF INVENTION: Of A
;   TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Paul W. Zimmerman
;   ADDRESSEE: Intellectual Property Services
;   ADDRESSEE: Battelle Memorial Institute
;   ADDRESSEE: PNNL P.O. Box 999
;   STREET: Washington Way
;   CITY: Richland
;   STATE: Washington
;   COUNTRY: U.S.A.
;   ZIP: 99352
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
;   COMPUTER: IBM PC/XT/AT
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: Word Processor (WordPerfect 5.1)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/853,659A
;   FILING DATE: Unknown
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: none
;   FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 28:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1797 bases
;   TYPE: nucleotide
;   STRANDEDNESS: double stranded
;   TOPOLOGY: linear
;   FEATURE:
;   OTHER INFORMATION: the coding nucleotides of SEQ ID
;   OTHER INFORMATION: NO:28 correspond to nucleotides 22630 through
;   OTHER INFORMATION: 24426 of SEQ ID NO:2
US-08-853-659A-28

```
  Query Match             8.1%;  Score 38.4;  DB 2;  Length 1797;
  Best Local Similarity  62.5%;  Pred. No. 0.36;
  Matches   60;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy        98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
             ||| ||||||||  |   ||||| |   | ||  |||  |||||     || || |  |
Db       229 AAATTGCCATAAAAAGAGATGAAAAAAAGTATTTAGTATAAAAAATCTTGTTCATCGATA 170

Qy       158 AACGTTTTTTCTTTATAGTATTTACTGATGCATTAT 193
             ||   |||||  | ||| | ||||||  |  | |||
Db       169 AAGAGTTTTTTATGATATTTTTTACTCTTATACTAT 134
```

RESULT 3
US-08-853-659A-6/c
; Sequence 6, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
;   APPLICANT: Wong, K.K.; Saffer, J.D.
;   TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
;   TITLE OF INVENTION: Of A
;   TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Paul W. Zimmerman
;   ADDRESSEE: Intellectual Property Services
;   ADDRESSEE: Battelle Memorial Institute
;   ADDRESSEE: PNNL P.O. Box 999
;   STREET: Washington Way
;   CITY: Richland
;   STATE: Washington
;   COUNTRY: U.S.A.
;   ZIP: 99352
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
;   COMPUTER: IBM PC/XT/AT
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: Word Processor (WordPerfect 5.1)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/853,659A
;   FILING DATE: Unknown
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: none
;   FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 8967 bases
;   TYPE: nucleotide
;   STRANDEDNESS: double stranded
;   TOPOLOGY: linear
;   FEATURE:
;   OTHER INFORMATION: SEQ ID NO:6 corresponds to
;   OTHER INFORMATION: nucleotides 15735 through 24701 of SEQ ID NO:2
US-08-853-659A-6

```
  Query Match             8.1%;  Score 38.4;  DB 2;  Length 8967;
  Best Local Similarity  62.5%;  Pred. No. 0.53;
  Matches   60;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy        98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
             ||| ||||||||  |   ||||| |   | ||  |||  |||||     || || |  |
Db      7124 AAATTGCCATAAAAAGAGATGAAAAAAAGTATTTAGTATAAAAAATCTTGTTCATCGATA 7065

Qy       158 AACGTTTTTTCTTTATAGTATTTACTGATGCATTAT 193
             ||   |||||  | ||| | ||||||  |  | |||
Db      7064 AAGAGTTTTTTATGATATTTTTTACTCTTATACTAT 7029
```

RESULT 4
US-08-853-659A-9
; Sequence 9, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
;   APPLICANT: Wong, K.K.; Saffer, J.D.
;   TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
;   TITLE OF INVENTION: Of A
;   TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
;   NUMBER OF SEQUENCES: 67
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Paul W. Zimmerman
;   ADDRESSEE: Intellectual Property Services
;   ADDRESSEE: Battelle Memorial Institute
;   ADDRESSEE: PNNL P.O. Box 999
;   STREET: Washington Way
;   CITY: Richland
;   STATE: Washington
;   COUNTRY: U.S.A.
;   ZIP: 99352
;   COMPUTER READABLE FORM:

```
;  CURRENT APPLICATION NUMBER: US/09/764,853
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 939
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 137
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-853-137

  Query Match             92.6%;  Score 438.8;  DB 9;  Length 530;
  Best Local Similarity   96.6%;  Pred. No. 6.9e-111;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |     |||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||:|||  |||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473


RESULT 5
US-09-764-877-986
; Sequence 986, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 986
;   LENGTH: 530
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-877-986

  Query Match             92.6%;  Score 438.8;  DB 10;  Length 530;
  Best Local Similarity   96.6%;  Pred. No. 6.9e-111;
  Matches  458;  Conservative    1;  Mismatches   13;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 61

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 121

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 181

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        182 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 241

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |     |||||||||||||
Db        242 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 301

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 361

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||:|||  |||||||||||||||||
Db        362 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAWAGG--TCAGAAAACATATGCTC 419

Qy        421 CATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        420 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 473


RESULT 6
US-09-989-442-33
; Sequence 33, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ08
;  CURRENT APPLICATION NUMBER: US/09/989,442
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/225,757
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/226,868
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/216,647
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,267
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/216,880
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,270
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/251,869
;  PRIOR FILING DATE: 2000-12-08
```

; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14

```
Qy        158 AACGTTTTTCTTTATAGTATTTACTGATGCATTAT 193
              ||   ||||| | ||| | |||||| |  | |||
Db      22798 AAGAGTTTTTATGATATTTTTACTCTTATACTAT 22763

RESULT 10
US-08-853-659A-61
; Sequence 61, Application US/08853659A
; Patent No. 5925522
;  GENERAL INFORMATION:
;    APPLICANT:  Wong, K.K.; Saffer, J.D.
;    TITLE OF INVENTION:  A Salmonella Sequence, Methods Of Detection
;    TITLE OF INVENTION:  Of A
;    TITLE OF INVENTION:  Salmonella Sequence, And Methods Of Detection Of Salmonella
;    NUMBER OF SEQUENCES:  67
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Paul W. Zimmerman
;      ADDRESSEE:  Intellectual Property Services
;      ADDRESSEE:  Battelle Memorial Institute
;      ADDRESSEE:  PNNL P.O. Box 999
;      STREET:  Washington Way
;      CITY:  Richland
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  99352
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette 3.50", 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Processor (WordPerfect 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/853,659A
;      FILING DATE:  Unknown
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  none
;      FILING DATE:  n/a
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24701 bases
;      TYPE:  nucleotide
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
US-08-853-659A-61

  Query Match              8.1%;  Score 38.4;  DB 2;  Length 24701;
  Best Local Similarity   38.5%;  Pred. No. 0.68;
  Matches   37;  Conservative   23;  Mismatches   36;  Indels    0;  Gaps    0;

Qy         98 AAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCA 157
              ||| :||||:|| |  |  :|||| |  | |:  ||:  |||||    :| :| : |  |
Db       1844 AAAUUGCCAUAAAAAGAGAUGAAAAAAAGUAUUUAGUAUAAAAAAUCUUGUUCAUCGAUA 1903

Qy        158 AACGTTTTTCTTTATAGTATTTACTGATGCATTAT 193
              ||   :::::  : |:| : :::||:  :   | :|:
Db       1904 AAGAGUUUUUUAUGAUAUUUUUUACUCUUAUACUAU 1939

RESULT 11
US-08-991-789A-169/c
; Sequence 169, Application US/08991789A
; Patent No. 6225054
;   GENERAL INFORMATION:
;        APPLICANT: Frudakis, Tony N.
;                   Smith, John M.
;                   Reed, Steven G.
;        TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;                            TREATMENT AND DIAGNOSIS OF BREAST CANCER
;        NUMBER OF SEQUENCES: 292
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed IP Law Group
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/991,789A
;             FILING DATE: 11-Dec-1997
;             CLASSIFICATION: &lt;Unknown&gt;
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Potter, Jane E. R.
;             REGISTRATION NUMBER: 33,332
;             REFERENCE/DOCKET NUMBER: 210121.419C3
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 169:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1265 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 169:
US-08-991-789A-169

  Query Match              8.0%;  Score 37.8;  DB 3;  Length 1265;
  Best Local Similarity   50.8%;  Pred. No. 0.47;
  Matches   90;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy        240 AGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGC 299
              | | | | ||||    || |   |||||| |            |||| |  |   ||||
Db        792 AAATAAAACAGGCTTCAGATTATCTTGGCTTTCATAATTATATTTTTCTTTTAAAGAAAA 733

Qy        300 TTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTT 359
                |     | ||  ||| || || |  |||       | | ||||||||| |  |  |  |    |  ||
Db        732 ATATCAACCCATTGTCAATGCACTGTTTTTCAAAGCATTTAAATAGAGGGTAAAACCCTT 673

Qy        360 TTGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATAT 416
              | |  || | |||||| ||    | || |  |      ||    || |  || ||||
Db        672 TGGAAATTAATACAGAAGAAATGATTCACTTTATGCATAAAAAATAAATAATAATAT 616

RESULT 12
US-09-062-451-169/c
; Sequence 169, Application US/09062451
; Patent No. 6344550
;  GENERAL INFORMATION:
;    APPLICANT:  Frudakis, Tony N.
;    APPLICANT:  Smith, John M.
;    APPLICANT:  Reed, Steven G.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;    TITLE OF INVENTION:  TREATMENT AND DIAGNOSIS OF BREAST CANCER
;    NUMBER OF SEQUENCES:  297
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/062,451
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  Maki, David J.
;       REGISTRATION NUMBER:  31,392
;       REFERENCE/DOCKET NUMBER:  210121.419C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  169:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1265 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-062-451-169

  Query Match             8.0%;  Score 37.8;  DB 4;  Length 1265;
  Best Local Similarity   50.8%;  Pred. No. 0.47;
  Matches   90;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy        240 AGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGC 299
              | | | | ||||   || |   |||||| |            |||| | |  ||||
Db        792 AAATAAAACAGGCTTCAGATTATCTTGGCTTTCATAATTATATTTTCTTTTAAAGAAAA 733

Qy        300 TTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTT 359
                |    | || ||| || || | |||         | | |||||||| | | | |  | |  ||
Db        732 ATATCAACCCATTGTCAATGCACTGTTTTCAAAGCATTTAAATAGAGGGTAAAACCCTT 673

Qy        360 TTGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATAT 416
              | |  || | ||||||  ||    | ||  |  |      ||    || |  ||  ||||
Db        672 TGGAAATTAATACAGAAGAAATGATTCACTTTATGCATAAAAAATAAATAATAATAT 616


RESULT 13
US-09-598-326-169/c
; Sequence 169, Application US/09598326
; Patent No. 6423496
;   GENERAL INFORMATION:
;        APPLICANT: Frudakis, Tony N.
;                   Smith, John M.
;                   Reed, Steven G.
;        TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;                            TREATMENT AND DIAGNOSIS OF BREAST CANCER
;        NUMBER OF SEQUENCES: 247
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group PLLC
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/598,326
;             FILING DATE: 20-Jun-2000
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Potter, Jane E.R.
;             REGISTRATION NUMBER: 33,332
;             REFERENCE/DOCKET NUMBER: 210121.419D1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 169:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1265 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 169:
US-09-598-326-169

  Query Match             8.0%;  Score 37.8;  DB 4;  Length 1265;
  Best Local Similarity   50.8%;  Pred. No. 0.47;
  Matches   90;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy        240 AGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGC 299
              | | | | ||||   || |   |||||| |            |||| | |  ||||
Db        792 AAATAAAACAGGCTTCAGATTATCTTGGCTTTCATAATTATATTTTCTTTTAAAGAAAA 733

Qy        300 TTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTT 359
                |    | || ||| || || | |||         | | |||||||| | | | |  | |  ||
Db        732 ATATCAACCCATTGTCAATGCACTGTTTTCAAAGCATTTAAATAGAGGGTAAAACCCTT 673

Qy        360 TTGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATAT 416
              | |  || | ||||||  ||    | ||  |  |      ||    || |  ||  ||||
Db        672 TGGAAATTAATACAGAAGAAATGATTCACTTTATGCATAAAAAATAAATAATAATAT 616


RESULT 14
US-09-289-198-169/c
; Sequence 169, Application US/09289198
; Patent No. 6586570
; GENERAL INFORMATION:
;  APPLICANT: Frudakis, Tony N.
;  APPLICANT: Smith, John M.
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Misher, Lynda
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.419C5
;  CURRENT APPLICATION NUMBER: US/09/289,198
;  CURRENT FILING DATE: 1999-04-09
;  EARLIER APPLICATION NUMBER: US 09/062,451
;  EARLIER FILING DATE: 1998-04-17
;  EARLIER APPLICATION NUMBER: US 08/991,789
;  EARLIER FILING DATE: 1997-12-11
;  EARLIER APPLICATION NUMBER: US 08/838,762
;  EARLIER FILING DATE: 1997-04-09
;  EARLIER APPLICATION NUMBER: PCT/US97/00485
;  EARLIER FILING DATE: 1997-01-10
;  EARLIER APPLICATION NUMBER: US 08/700,014
;  EARLIER FILING DATE: 1996-08-20
;  EARLIER APPLICATION NUMBER: US 08/585,392
;  EARLIER FILING DATE: 1996-01-01
;  NUMBER OF SEQ ID NOS: 312
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 169
;   LENGTH: 1265
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-289-198-169

  Query Match             8.0%;  Score 37.8;  DB 4;  Length 1265;
  Best Local Similarity   50.8%;  Pred. No. 0.47;
  Matches   90;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy        240 AGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGC 299
              | | | | ||||   || |   |||||| |            |||| | |  ||||
Db        792 AAATAAAACAGGCTTCAGATTATCTTGGCTTTCATAATTATATTTTCTTTTAAAGAAAA 733

Qy        300 TTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTT 359
                |    | || ||| || || | |||         | | |||||||| | | | |  | |  ||
Db        732 ATATCAACCCATTGTCAATGCACTGTTTTCAAAGCATTTAAATAGAGGGTAAAACCCTT 673

Qy        360 TTGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATAT 416
              | |  || | ||||||  ||    | ||  |  |      ||    || |  ||  ||||
Db        672 TGGAAATTAATACAGAAGAAATGATTCACTTTATGCATAAAAAATAAATAATAATAT 616


RESULT 15
US-09-134-001C-1232/c
; Sequence 1232, Application US/09134001C
; Patent No. 6380370
```

```
;  NUMBER OF SEQ ID NOS: 305
;  SOFTWARE: PERL Program
; SEQ ID NO 255
;   LENGTH: 1126
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030120057A1 3556218CT1
US-10-313-542-255

  Query Match             92.3%;  Score 437.6;  DB 14;  Length 1126;
  Best Local Similarity   96.6%;  Pred. No. 2.2e-110;
  Matches  458;  Conservative    0;  Mismatches   14;  Indels    2;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCACA 173

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTCATCAAAGTGCCATAACAGCAACTGAA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 233

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 293

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||
Db        294 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 353

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              |||||||||||||||||||||||||||||||||| | | ||||||||||||||||||||
Db        354 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTATTCTGCCCATTAACAGTGCT 413

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        414 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 473

Qy        361 TGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGATCAGAAAACATATGCTC 420
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 TGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTC 531

Qy        421 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 585


RESULT 12
US-09-895-686-37/c
; Sequence 37, Application US/09895686
; Patent No. US20020106655A1
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT: Lal, Preeti
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Baughn, Mariah R.
;  TITLE OF INVENTION: HUMAN GPCR PROTEINS
;  FILE REFERENCE: PC-0044 CIP
;  CURRENT APPLICATION NUMBER: US/09/895,686
;  CURRENT FILING DATE:  2001-06-28
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PERL Program
; SEQ ID NO 37
;   LENGTH: 612
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020106655A1 2488822X308B1
;   NAME/KEY: unsure
;   LOCATION: 561
;   OTHER INFORMATION: a, t, c, g, or other

US-09-895-686-37

  Query Match             92.1%;  Score 436.6;  DB 10;  Length 612;
  Best Local Similarity   96.6%;  Pred. No. 3e-110;
  Matches  457;  Conservative    0;  Mismatches   14;  Indels    2;  Gaps    1;

Qy          2 CCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAG 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 CCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCACAG 438

Qy         62 TTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAA 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 TTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAA 378

Qy        122 TACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTTA 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 TACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTA 318

Qy        182 CTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAG 241
              |||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db        317 CTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTAG 258

Qy        242 AAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTT 301
              |||||||||||||||||||||||||||||||| | | ||||||||||||||||||||||
Db        257 AAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTATTCTGCCCATTAACAGTGCTT 198

Qy        302 TGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTT 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 TGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTT 138

Qy        362 GGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTATCAGAAAACATATGCTCC 421
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGG--TCAGAAAACATATGCTCC 80

Qy        422 ATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 ATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 27


RESULT 13
US-10-321-807-17
; Sequence 17, Application US/10321807
; Publication No. US20030166148A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Rupong
;  APPLICANT: Dang, Huong T.
;  APPLICANT: Lowitz, Kevin P.
;  TITLE OF INVENTION: No. US20030166148A1-Endogenous, Constitutively Activated Huma
;  TITLE OF INVENTION: Receptors
;  FILE REFERENCE: ARENO086
;  CURRENT APPLICATION NUMBER: US/10/321,807
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: US/09/714,008
;  PRIOR FILING DATE: 2000-11-16
;  PRIOR APPLICATION NUMBER: 09/170,496
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: PCT/US99/23938
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: 60/166,088
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/166,099
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/166,369
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/171,902
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: 60/171,901
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: 60/171,900
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: 60/181,749
;  PRIOR FILING DATE: 2000-02-11
;  Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
;  NUMBER OF SEQ ID NOS: 133
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-321-807-17

  Query Match             36.2%;  Score 171.8;  DB 12;  Length 1068;
  Best Local Similarity   67.4%;  Pred. No. 4.1e-37;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
              |||||  | | |||||| || ||| | || |||     || ||| ||||||| || ||||||
Db        568 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 627

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
              |||   |  ||||| | |     ||||  |||  ||| |       | || ||| || |||||
Db        628 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 687

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTACTGATGCA 189
                   ||  || |||| ||     |||| || ||||| || |||||||| ||  |||||||
Db        688 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 747

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
                | |||||||| ||   | |||||||  |||     ||||| ||   | | ||| |||||||||
Db        748 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 807

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
              |  || || || || ||| ||||||||  ||      | |  ||||||||||||||| |||
Db        808 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 867

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
              || ||||||||||| || || |       ||||| || | | |  | | | |    ||| ||
Db        868 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 926
```

```
RESULT 14
US-09-928-175-11
; Sequence 11, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618Alma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 2049
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2046)
;   NAME/KEY: sig_peptide
;   LOCATION: (1)..(108)
US-09-928-175-11

  Query Match             36.2%;  Score 171.8;  DB 10;  Length 2049;
  Best Local Similarity   67.4%;  Pred. No. 5.8e-37;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
              |||||  | | |||||| || ||| | || |||     || ||| ||||||| || ||||||
Db       1549 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 1608

Qy         70 IATGGAAGCATGTTTTATAGTGTTCAICAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
              |||   |  ||||| | |     ||||  |||  ||| |       | || ||| || |||||
Db       1609 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 1668

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTACTGATGCA 189
                   ||  || |||| ||     |||| || ||||| || |||||||| ||  |||||||
Db       1669 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCTCTGATGCC 1728

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
                | |||||||| ||   | |||||||  |||     ||||| ||   | | ||| |||||||||
Db       1729 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 1788

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
              |  || || || || ||| ||||||||  ||      | |  ||||||||||||||| |||
Db       1789 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 1848

Qy        310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
              || ||||||||||| || || |       ||||| || | | |  | | | |    ||| ||
Db       1849 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 1907
```

```
RESULT 15
US-09-965-536A-5
; Sequence 5, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
;  APPLICANT: FEDER, J. N.
;  APPLICANT:  MINTIER, G.
;  APPLICANT:  RAMANATHAN, C. S.
;  APPLICANT:  HAWKEN, D. R.
;  TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
;  TITLE OF INVENTION:  EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
;  FILE REFERENCE: DO041NP
;  CURRENT APPLICATION NUMBER: US/09/965,536A
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: 60/235,713
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/261,781
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: 60/306,605
;  PRIOR FILING DATE: 2001-07-19
;  PRIOR APPLICATION NUMBER: 60/310,436
;  PRIOR FILING DATE: 2001-08-03
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 2142
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-965-536A-5

  Query Match             36.2%;  Score 171.8;  DB 11;  Length 2142;
  Best Local Similarity   67.4%;  Pred. No. 5.9e-37;
  Matches  242;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         10 TATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC 69
              |||||  | | |||||| || ||| | || |||     || ||| ||||||| || ||||||
Db       1642 TATTCTCTTGGAATTTTCCTAGGTGTGAACTTGCTGGCTTTTCTCATCATTGTGTTTTCC 1701

Qy         70 TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAI 129
              |||   |  ||||| | |     ||||  |||  ||| |       | || ||| || |||||
Db       1702 TATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAGTAAGGAAT 1761

Qy        130 CAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTACTGATGCA 189
                   ||  || |||| ||     |||| || ||||| || |||||||| ||  |||||||
Db       1762 TGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTTATAGTGTTCICTGATGCC 1821

Qy        190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
                | |||||||| ||   | |||||||  |||     ||||| ||   | | ||| |||||||||
Db       1822 ATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGGAAATACCA 1881

Qy        250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
```

```
             |  || || || || ||| ||||||||  ||    | |  |||||||||||||||| |||
Db      1882 GACACAATGACTTCCTGGATAGTGATTTTTTTCCTTCCAGTTAACAGTGCTTTGAATCCA 1941

Qy       310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAA 368
             || ||||||||||| || || |     ||||| || | | |  | | | |   ||| ||
Db      1942 ATCCTCTATACTCTCACAACCAACTTTTTTAAGGACAAGTTGAAACAGCTGCTGCACAA 2000
```

Search completed: September 25, 2003, 15:11:48
Job time : 191 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:52:51 ; Search time 41 Seconds
(without alignments)
607.807 Million cell updates/sec

Title: US-10-049-568-2
Perfect score: 826
Sequence: 1 AQIYSVAIPLGINLAAFIII.....HMLHESSGGKCGHCRRCHLS 157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	826	100.0	157	22	AAB35407	Human HGRL101 G-pr
2	638	77.2	188	22	AAM99945	Human expressed po
3	638	77.2	209	22	AAU87243	Novel central nerv
4	638	77.2	396	21	AAY57286	Human GPCR protein
5	638	77.2	396	22	AAD02498	Human CON222 G pro
6	638	77.2	396	23	ABU54868	Human G-protein co
7	638	77.2	722	20	AAY42171	Human LGR7 short f
8	638	77.2	757	20	AAY42170	Human LGR7 long fo
9	638	77.2	757	24	ABP81724	Human G protein-co
10	632	76.5	176	22	AAU87545	Novel central nerv
11	632	76.5	176	22	AAM99972	Human expressed po
12	632	76.5	176	22	ABB04062	Human musculoskele
13	632	76.5	176	22	ABB10249	Human cDNA SEQ ID
14	632	76.5	176	22	AAU18108	Novel human uterin
15	632	76.5	176	22	AAU18344	Human endocrine po
16	632	76.5	176	22	AAU18666	Renal and cardiova
17	632	76.5	176	22	AAU21646	Novel human peopla
18	632	76.5	176	23	ABP66836	Human polypeptide
19	632	76.5	176	23	ABJ05735	Novel human prost
20	632	76.5	176	24	ABU13356	Novel human muscul
21	586.5	71.0	140	21	AAB41526	Human ORFX ORF1290
22	392.5	47.5	355	22	AAU04370	Human G-protein co
23	392.5	47.5	355	23	ABB06253	Human G protein-co
24	392.5	47.5	491	23	ABB06252	Human G protein-co
25	392.5	47.5	515	23	ABB06251	Human G protein-co
26	392.5	47.5	610	23	ABB06250	Human G protein-co
27	392.5	47.5	713	23	AAD97159	Human G-protein co
28	392.5	47.5	726	23	AAB71322	Human GCREC-1 ENCY
29	392.5	47.5	730	23	ABB06255	Human G protein-co
30	392.5	47.5	737	23	AAD97158	Human G protein co
31	392.5	47.5	754	23	ABB06254	Human G protein-co
32	392.5	47.5	754	24	AAO16303	Human LGR9 G-prote
33	391.5	47.4	363	24	ABU37884	NOVX protein seque
34	351	42.5	108	23	AAE20167	Human nGPCR-1079 v
35	349	42.3	107	23	AAE20148	Human novel G-prot
36	349	42.3	109	23	AAE20168	Human nGPCR-1079 v
37	344	29.5	327	22	AAG80961	Human nGPCR57. Ho
38	244	29.5	327	23	ABG93779	Human G protein-co
39	224.5	27.2	334	22	ABB60463	Drosophila melanog
40	224.5	27.2	334	22	AAU38929	Drosophila G-prote
41	212.5	25.7	359	22	ABB61216	Drosophila melanog
42	212.5	25.7	359	22	AAU38931	Drosophila G-prote
43	207.5	25.1	1115	23	AAM27959	Lymnaea stagnalis
44	139.5	16.9	620	14	AAR30522	N-terminal of LH r
45	139.5	16.9	634	14	AAR30520	N-terminal of LH r

ALIGNMENTS

RESULT 1
AAB35407
ID AAB35407 standard; Protein; 157 AA.
XX
AC AAB35407;
XX
DT 23-MAY-2001 (first entry)
XX
DE Human HGRL101 G-protein coupled receptor.
XX
KW Human; HGRL101; G-protein coupled receptor; infection; pain; cancer;
KW diabetes; obesity; eating disorder; asthma; Parkinson's disease;
KW hypotension; osteoporosis; myocardial infarction; migraine; allergy;
KW psychotic disorder; neurological disorder; dyskinesia; vaccine.
XX
OS Homo sapiens.
XX
PN WO200114548-A2.
XX
PD 01-MAR-2001.
XX
PF 09-AUG-2000; 2000WO-EP07723.
XX
PR 19-AUG-1999; 99EP-0116345.
XX
PA (MERR) MERCK PATENT GMBH.
XX
PI Duecker K;
XX
DR WPI; 2001-226617/23.
DR N-PSDB; AAF28059.
XX

```
PT   Novel G-coupled protein receptor, HGRL101 useful for treating diseases
PT   such as microbial infections, cancers, obesity, asthma, diabetes,
PT   hypotension, osteoporosis, myocardial infarction, stroke, ulcer,
PT   allergy -
XX
PS   Claim 1; Page 36; 36pp; English.
XX
CC   The present invention provides the protein and coding sequences for a
CC   novel human G-protein coupled receptor, designated HGRL101. The sequences
CC   are useful in the diagnosis, prevention and treatment of diseases
CC   including infections, pain, cancer, diabetes, anorexia, bulimia, asthma,
CC   Parkinson's disease, acute heart failure, hypotension, hypertension,
CC   urinary retention, osteoporosis, angina pectoris, myocardial infarction,
CC   stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
CC   vomiting, psychotic and neurological disorders such as anxiety,
CC   schizophrenia, manic depression, depression, delirium, dementia and
CC   severe mental retardation, and dyskinesias including Huntington's disease
CC   and Gilles dela Tourette's syndrome. The present sequence is the HGRL101
CC   protein.
XX
SQ   Sequence   157 AA;

  Query Match             100.0%;  Score 826;  DB 22;  Length 157;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-88;
  Matches  157;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRMQVKKEMILAKRFFFIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRMQVKKEMILAKRFFFIVF 60

Qy         61 TDALCWIPIPVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFKEMIHRF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TDALCWIPIPVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFKEMIHRF 120

Qy        121 WNHYRQRKSMDSKGIRKEMLHHSSGGKCGHCRRCHLS 157
              |||||||||||||||||||||||||||||||||||||
Db        121 WNHYRQRKSMDSKGIRKEMLHHSSGGKCGHCRRCHLS 157


RESULT 2
AAM99945
ID   AAM99945 standard; Protein; 188 AA.
XX
AC   AAM99945;
XX
DT   04-JAN-2002  (first entry)
XX
DE   Human expressed polypeptide SEQ ID NO 69.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200155387-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01310.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
```

```
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244517.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259878.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-465573/50.
DR   N-PSDB; AAI99557.
XX
PT   Isolated digestive system associated polypeptide for treating,
PT   preventing and/ or prognosing disorders related to the digestive system
PT   including digestive system cancers and also for testing and detection
PT   e.g. diagnosis -
XX
PS   Claim 11; SEQ ID NO 69; 509pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC   (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (anti)agonists are useful
CC   in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC   (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   188 AA;

  Query Match          77.2%;  Score 638;  DB 22;  Length 188;
  Best Local Similarity   94.8%;  Pred. No. 1.8e-66;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFSVHQSAITAFEIRNQVKKEMILAKRFFIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 AQIYSVAIFLGINLAAFIIIVFSYGSMFSVHQSAITAFEIRNQVKKEMILAKRFFIVF 66

Qy         61 TDALCWIPIFVVAKPLSLLQVEIPGTTSWVVIGYSAINSALNPILYTLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||
Db         67 TDALCWIPIFVVKPLSLLQVEIPGTITSWVVIPILPINSALNPILYTLTRPFKEMIHRF 126

Qy        121 WHNYRQRKSMDSKG 134
               |:|||||||||||
Db        127 WYNYRQRKSMDSKG 140


RESULT 3
AAU87243
ID   AAU87243 standard; Protein; 209 AA.
XX
AC   AAU87243;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Novel central nervous system protein #153.
XX
KW   Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW   hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW   cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW   nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW   acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW   adenocarcinoma; reproductive system disorder; testicular feminisation;
KW   endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW   respiratory disorder; renal disorder; kidney failure; blood disorder;
KW   myocardial infarction; wound healing; cell proliferation; skin aging;
KW   food additive; food preservative; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200155318-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01332.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
```

PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215335.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-581633/65.
DR N-PSDB; ABX43573.
XX
PT New isolated nucleic acid encoding a protein for diagnosing,
PT preventing, treating or ameliorating medical conditions and used as
PT food additives or preservatives "
XX
PS Claim 9; SEQ ID NO 761; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative

```
CC   disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC   e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC   angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC   amyliotrophic lateral sclerosis, infections caused by bacteria, viruses
CC   e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC   e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC   adenocarcinonas and irritable bowel syndrone, reproductive system
CC   disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC   and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC   leukaemia, disorders involving neovascularisation e.g. malignancies,
CC   respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC   acute kidney failure and blood related disorders e.g. myocardial
CC   infarction. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities, fat content, lipid, protein,

  Query Match             77.2%;  Score 638;  DB 22;  Length 209;
  Best Local Similarity   94.8%;  Pred. No. 2e-66;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIPLGINLAAPIIIVFSYGSMPYSVHQSAITATEIRNQVKKEMILAKRPFPIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 AQIYSVAIPLGINLAAPIIIVFSYGSMPYSVHQSAITATEIRNQVKKEMILAKRPFPIVF 87

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPLLYTLLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    |||||||||||||||||||||||||||
Db         88 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIPILPINSALNPILYTLLTTRPFKEMIHRF 147

Qy        121 WHNYRQRKSMDSKG 134
              |:||||||||||||
Db        148 WYNYRQRKSMDSKG 161



RESULT 4
AAY57286
ID   AAY57286 standard; Protein; 396 AA.
XX
AC   AAY57286;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Human GPCR protein (HGPRP) sequence (clone ID 2488822).
XX
KW   Human; G protein coupled protein receptor; HGPRP; cell proliferation;
KW   neurological; immune disorder; cytostatic; anti-arteriosclerotic;
KW   anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;
KW   immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
KW   anti-Alzheimer's; anti-Parkinsonian; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200015793-A2.
XX
PD   23-MAR-2000.
XX
PF   17-SEP-1999;  99WO-US20958.
XX
PR   17-SEP-1998;  98US-0156513.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Bandman O,  Lal P,  Tang YT,  Corley NC,  Guegler KJ,  Gorgone GA;
PI   Baughn MR;
XX
DR   WPI; 2000-271432/23.
DR   N-PSDB; AAZ90524.
XX
PT   Human G protein coupled protein receptor peptides useful for the
PT   prevention, diagnosis and treatment of cell proliferative, neurological
PT   and immune disorders -
XX
PS   Claim 1; Page 62-63; 71pp; English.
XX
CC   The invention provides human G protein coupled protein receptor (HGPRP)
CC   polypeptides and polynucleotides encoding them. The polypeptides can be
CC   produced by standard recombinant methodology. The polynucleotides and
CC   polypeptides may be used in the prevention, treatment and diagnosis of
CC   diseases associated with their inappropriate expression. Diseases that
CC   can be treated are cell proliferative disorders (e.g. arteriosclerosis,
CC   atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
CC   adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
CC   disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
CC   and Parkinson's disease). The anti-HGPCR antibodies may also be used as
CC   diagnostic agents for detecting the presence of HGPRP polypeptides in
CC   samples (e.g. by enzyme linked immunosorbant assay (ELISA)). Sequences
CC   AAY57283-288 represent the HGPRP polypeptides.
XX
SQ   Sequence  396 AA;

  Query Match             77.2%;  Score 638;  DB 21;  Length 396;
  Best Local Similarity   94.8%;  Pred. No. 4.7e-66;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIPLGINLAAPIIIVFSYGSMPYSVHQSAITATEIRNQVKKEMILAKRPFPIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 AQIYSVAIPLGINLAAPIIIVFSYGSMPYSVHQSAITATEIRNQVKKEMILAKRPFPIVF 274

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    |||||||||||||||||||||||||||
Db        275 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIPILPINSALNPILYTLLTTRPFKEMIHRF 334

Qy        121 WHNYRQRKSMDSKG 134
              |:||||||||||||
Db        335 WYNYRQRKSMDSKG 348



RESULT 5
AAE02498
ID   AAE02498 standard; Protein; 396 AA.
XX
AC   AAE02498;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human CON222 G protein-coupled receptor protein.
XX
KW   Human; G protein-coupled receptor; GPCR; CON222 protein; schizophrenia;
KW   neuroleptic; serotropic; neuroprotective; bipolar disease; psychotropic;
KW   neurological disorder; psychiatric disease; neurosis; anxiety; neuritis;
KW   attention deficit hyperactivity disorder; neurasthenia; senile dementia;
KW   affective disorder; neuropathy; Alzheimer's disease; Parkinson's disease;
KW   depression; migraine; genetic screening.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          42..65
FT                   /label= Transmembrane_domain_(1TM)
FT   Domain          66..78
FT                   /label= Intracellular_domain
FT                   /note= "First IC loop"
FT   Domain          79..103
FT                   /label= Transmembrane_domain_(2TM)
FT   Domain          104..124
FT                   /label= Extracellular_domain
FT                   /note= "First EC loop"
FT   Domain          125..156
FT                   /label= Transmembrane_domain_(3TM)
FT   Domain          157..166
FT                   /label= Intracellular_domain
FT                   /note= "Second IC loop"
FT   Domain          167..188
FT                   /label= Transmembrane_domain_(4TM)
```

```
FT   Domain          189..216
FT                   /label= Extracellular_domain
FT                   /note= "Second EC loop"
FT   Domain          217..241
FT                   /label= Transmembrane_domain_(5TM)
FT   Domain          242..267
FT                   /label= Intracellular_domain
FT                   /note= "Third IC loop"
FT   Domain          268..290
FT                   /label= Transmembrane_domain_(6TM)
FT   Domain          291..300
FT                   /label= Extracellular_domain
FT                   /note= "Third EC loop"
FT   Domain          301..320
FT                   /label= Transmembrane_domain_(7TM)
XX
PN   WO200131014-A2.
XX
PD   03-MAY-2001.
XX
PF   27-OCT-2000; 2000WO-US29601.
XX
PR   27-OCT-1999;   99US-0427653.
PR   27-OCT-1999;   99US-0427859.
PR   27-OCT-1999;   99US-0428020.
PR   27-OCT-1999;   99US-0428114.
PR   28-OCT-1999;   99US-0429517.
PR   28-OCT-1999;   99US-0429555.
PR   28-OCT-1999;   99US-0429676.
PR   28-OCT-1999;   99US-0429695.
PR   03-DEC-1999;   99US-0454399.
PR   12-JAN-2000; 2000US-0481794.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Vogeli G,  Wood LS,  Merchant K;
XX
DR   WPI; 2001-328653/34.
DR   N-PSDB; AAD06507.
XX
PT   Seven transmembrane receptor polypeptides and polynucleotides, useful
PT   for treating neurological or psychiatric disorders, e.g. schizophrenia,
PT   as well as for identifying compounds useful for treating schizophrenia
PT   -
XX
PS   Claim 1; Page 15-16; 215pp; English.
XX
CC   The invention relates to human G protein-coupled receptor (GPCR) and
CC   their corresponding DNA molecules. GPCR is also referred as seven
CC   transmembrane receptor. G protein-coupled receptor protein is useful for
CC   treating neurological disorder, particularly schizophrenia. GPCR protein
CC   is also useful for identifying compounds useful for treating
CC   schizophrenia. These compounds are also useful for treating other
CC   neurological and psychiatric diseases, e.g. depression, anxiety, bipolar
CC   disease, affective disorders, attention deficit hyperactivity disorder/
CC   attention deficit disorder, epilepsy, neuritis, neurasthenia, neuropathy,
CC   neurosis, Alzheimer's disease, Parkinson's disease, migraine and senile
CC   dementia. The invention also provides genetic screening procedures that
CC   entail analysing a person's genome with respect to GPCR. The vectors are
CC   useful for the recombinant production of the GPCR's. The present sequence
CC   is human CON222 G protein-coupled receptor (GPCR) protein.
XX
SQ   Sequence   396 AA;

  Query Match            77.2%;  Score 638;  DB 22;  Length 396;
  Best Local Similarity  94.8%;  Pred. No. 4.7e-66;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 274

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
                |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||
Db        275 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF 334

Qy        121 WHNYRQRKSMDSKG 134
                |:||||||||||||
Db        335 WYNYRQRKSMDSKG 348


RESULT 6
ABU54868
ID   ABU54868 standard; Protein; 396 AA.
XX
AC   ABU54868;
XX
DT   14-MAR-2003  (first entry)
XX
DE   Human G-protein coupled receptor #4, Incyte clone 2488822CD1.
XX
KW   Human; G-protein coupled receptor; GPCR; neoplastic disorder;
KW   neurological disorder; immune disorder; cytostatic; pancreatic cancer;
KW   follicular carcinoma of the thyroid; leiomyoma of the uterus; epilepsy;
KW   interstitial nephritis; cancer.
XX
OS   Homo sapiens.
XX
PN   US2002106655-A1.
XX
PD   08-AUG-2002.
XX
PF   28-JUN-2001; 2001US-0895686.
XX
PR   17-SEP-1998;   98US-0156513.
XX
PA   (BAND/) BANDMAN O.
PA   (LALP/) LAL P G.
PA   (TANG/) TANG Y T.
PA   (BAUG/) BAUGHN M R.
XX
PI   Bandman O,  Lal PG,  Tang YT,  Baughn MR;
XX
DR   WPI; 2002-697866/75.
DR   N-PSDB; ABX73054.
XX
PT   New cDNAs encoding G protein coupled receptors are useful for the
PT   diagnosis, prognosis, treatment and evaluation of therapies for
PT   neoplastic, neurological and immune disorders  -
XX
PS   Claim 13; Fig 2; 61pp; English.
XX
CC   The invention relates to an isolated cDNA encoding G-protein
CC   coupled receptor (GPCR) appearing as ABU54865-ABU54870.
CC   Also included are fragments of the cDNAs, species variants having
CC   at least 75% identity to the cDNAs, vectors comprising the cDNAs,
CC   a host cells comprising the above vectors, producing a protein
CC   (comprising culturing the above host cell under expression conditions and
CC   recovering the protein), using a cDNA to detect expression of a
CC   nucleic acid in a sample or to screen for compounds or molecules
CC   which bind to the cDNAs, using the GPCR proteins to screen
CC   compounds or molecules for ligands, using a GPCR protein to prepare and
CC   purify antibodies, an anti-GPCR antibody and using the antibody to detect
CC   expression of a GPCR protein in a sample and is diagnostic of cancer.
CC   The invention is useful for the diagnosis, prognosis, treatment
CC   and evaluation of therapies for neoplastic, neurological and immune
CC   disorders, particularly follicular carcinoma of the thyroid, leiomyoma
CC   of the uterus, pancreatic cancer, epilepsy, interstitial nephritis and
CC   immune response as a complication of cancer. The present sequence
CC   is a human GPCR protein of the invention.
XX
SQ   Sequence   396 AA;

  Query Match            77.2%;  Score 638;  DB 23;  Length 396;
  Best Local Similarity  94.8%;  Pred. No. 4.7e-66;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 274

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||||
Db        275 TDALCWIPIFVVKPLSLLQVEIPGTITSWVVIPLLPINSALNPILYTLTTRPFKEMIHRF 334

Qy        121 WHNYRQRKSMDSKG 134
              |:||||||||||||
Db        335 WYNYRQRKSMDSKG 348


RESULT 7
AAY42171
ID   AAY42171 standard; Protein; 722 AA.
XX
AC   AAY42171;
XX
DT   20-DEC-1999  (first entry)
XX
DE   Human LGR7 short form protein sequence.
XX
KW   Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
KW   extracellular leucine rich repeat region; mapping; identification.
XX
OS   Homo sapiens.
XX
PN   WO9948921-A1.
XX
PD   30-SEP-1999.
XX
PF   25-MAR-1999;   99WO-US06573.
XX
PR   26-MAR-1998;   98US-0079501.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PA   (ORGA ) ORGANON NV.
XX
PI   Hsueh AJW,  Hsu SY,  Liang S,  Van Der Spek PJ;
XX
DR   WPI; 1999-591074/50.
DR   N-PSDB; AAZ25346.
XX
PT   New G-protein coupled receptors, useful for identifying their own
PT   ligands -
XX
PS   Claim 2; Fig 4; 54pp; English.
XX
CC   The present sequence represents the human G-protein coupled receptor
CC   having extracellular leucine rich repeat regions, designated LGR7 short
CC   form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
CC   the receptor. The polypeptides and/or polynucleotides are also useful
CC   for homologous or related genes, producing compositions that modulate
CC   the expression or function of the receptors, gene therapy, mapping
CC   functional regions of the receptors, studying associated physiological
CC   pathways, in vivo prophylactic and therapeutic purposes, as immunogens
CC   for producing antibodies, and for identifying biologically active
CC   agents. The polypeptides contain a G-protein coupled seven
CC   transmembrane region and a leucine rich repeat extracellular domain.
CC   These regions capture and facilitate optimal orientation of its ligand.
CC   The proteins are also expressed in diverse tissues.
XX
SQ   Sequence   722 AA;

  Query Match           77.2%;  Score 638;  DB 20;  Length 722;
  Best Local Similarity   94.8%;  Pred. No. 1e-65;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 600

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||||
Db        601 TDALCWIPIFVVKPLSLLQVEIPGTITSWVVIPLLPINSALNPILYTLTTRPFKEMIHRF 660

Qy        121 WHNYRQRKSMDSKG 134
              |:||||||||||||
Db        661 WYNYRQRKSMDSKG 674


RESULT 8
AAY42170
ID   AAY42170 standard; Protein; 757 AA.
XX
AC   AAY42170;
XX
DT   20-DEC-1999  (first entry)
XX
DE   Human LGR7 long form protein sequence.
XX
KW   Human; LGR4; LGR5; LGR7; G-protein coupled receptor; gene therapy;
KW   extracellular leucine rich repeat region; mapping; identification.
XX
OS   Homo sapiens.
XX
PN   WO9948921-A1.
XX
PD   30-SEP-1999.
XX
PF   25-MAR-1999;   99WO-US06573.
XX
PR   26-MAR-1998;   98US-0079501.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
PA   (ORGA ) ORGANON NV.
XX
PI   Hsueh AJW,  Hsu SY,  Liang S,  Van Der Spek PJ;
XX
DR   WPI; 1999-591074/50.
DR   N-PSDB; AAZ25345.
XX
PT   New G-protein coupled receptors, useful for identifying their own
PT   ligands -
XX
PS   Claim 2; Fig 3; 54pp; English.
XX
CC   The present sequence represents the human G-protein coupled receptor
CC   having extracellular leucine rich repeat regions, designated LGR7 long
CC   form. The LGR4, LGR5 and LGR7 proteins are used to identify ligands for
CC   the receptor. The polypeptides and/or polynucleotides are also useful
CC   for homologous or related genes, producing compositions that modulate
CC   the expression or function of the receptors, gene therapy, mapping
CC   functional regions of the receptors, studying associated physiological
CC   pathways, in vivo prophylactic and therapeutic purposes, as immunogens
CC   for producing antibodies, and for identifying biologically active
CC   agents. The polypeptides contain a G-protein coupled seven
CC   transmembrane region and a leucine rich repeat extracellular domain.
CC   These regions capture and facilitate optimal orientation of its ligand.
CC   The proteins are also expressed in diverse tissues.
XX
SQ   Sequence   757 AA;

  Query Match           77.2%;  Score 638;  DB 20;  Length 757;
  Best Local Similarity   94.8%;  Pred. No. 1.1e-65;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 635

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||||
Db        636 TDALCWIPIFVVKPLSLLQVEIPGTITSWVVIPLLPINSALNPILYTLTTRPFKEMIHRF 695
```

```
Qy        121 WHNYRQRKSMDSKG 134
              |:|||||||||||||
Db        696 WYNYRQRKSMDSKG 709

RESULT 9
ABP81724
ID   ABP81724 standard; Protein; 757 AA.
XX
AC   ABP81724;
XX
DT   04-MAR-2003  (first entry)
XX
DE   Human G protein-coupled receptor LGR7 protein SEQ ID NO:623.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
KW   Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW   osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW   graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW   psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW   mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW   hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW   ulcer.
XX
OS   Homo sapiens.
XX
PN   WO200261087-A2.
XX
PD   08-AUG-2002.
XX
PF   19-DEC-2001; 2001WO-US50107.
XX
PR   19-DEC-2000; 2000US-257144P.
XX
PA   (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI   Burmer GC,  Roush CL,  Brown JP;
XX
DR   WPI; 2003-046718/04.
DR   N-PSDB; ABZ42570.
XX
PT   New isolated antigenic peptides e.g., for G protein-coupled receptors
PT   (GPCR), useful for diagnosing and designing drugs for treating
PT   conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
PT   cancer or autoimmune diseases  -
XX
PS   Disclosure; Fig 1; 523pp; English.
XX
CC   The present invention describes antigenic peptides (I) comprising:
CC   (a) any one of 1601 sequences (see ABP82019 to ABP83519) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular
CC   G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity
CC   or avidity for a particular GPCR. (I) can be used as GPCR modulators and
CC   in gene therapy. The antigenic peptides for GPCRs are useful in detecting
CC   an antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-relied disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC   GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   757 AA;

  Query Match          77.2%;  Score 638;  DB 24;  Length 757;
  Best Local Similarity   94.8%;  Pred. No. 1.1e-65;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIVFSYGSMPTSVHQSAITAFEIRNQVKKEMILAKRPFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 AQIYSVAIFLGINLAAFIIVFSYGSMPTSVHQSAITAFEIRNQVKKEMILAKRPFFIVF 635

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    |||||||||||||||||||||||||
Db        636 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFLLPINSALNPILYTLTRPFKEMIHRF 695

Qy        121 WHNYRQRKSMDSKG 134
              |:|||||||||||
Db        696 WYNYRQRKSMDSKG 709

RESULT 10
AAU87545
ID   AAU87545 standard; Protein; 176 AA.
XX
AC   AAU87545;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Novel central nervous system protein #455.
XX
KW   Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW   hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW   cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW   nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW   acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW   adenocarcinoma; reproductive system disorder; testicular feminisation;
KW   endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW   respiratory disorder; renal disorder; kidney failure; blood disorder;
KW   myocardial infarction; wound healing; cell proliferation; skin aging;
KW   food additive; food preservative; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200155318-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01332.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
```

PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226858.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-581633/65.
DR N-PSDB; ABK43875.
XX
PT New isolated nucleic acid encoding a protein for diagnosing,
PT preventing, treating or ameliorating medical conditions and used as
PT food additives or preservatives -
XX
PS Claim 9; SEQ ID NO 1063; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (II) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyliotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome, reproductive system
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.

```
CC   acute kidney failure and blood related disorders e.g. myocardial
CC   infarction. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities, fat content, lipid, protein,

  Query Match          76.5%;  Score 632;  DB 22;  Length 176;
  Best Local Similarity   94.0%;  Pred. No. 8.3e-66;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIPLGINLAAPIIIVFSYGSMFYSVHQSAITAFEIRNQVKKEMILAKRFFPIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIPLGINLAAPIIIVFSYGSMFYSVHQSAITAFEIRNQVKKEMILAKRFFPIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||    ||||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKPLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSKG 134
              |:|||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 11
AAM99972
ID   AAM99972 standard; Protein; 176 AA.
XX
AC   AAM99972;
XX
DT   04-JAN-2002  (first entry)
XX
DE   Human expressed polypeptide SEQ ID NO 96.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200155387-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01310.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184654.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235634.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236366.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
```

PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250191.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465573/50.
DR N-PSDB; AAI99584.
XX
PT Isolated digestive system associated polypeptide for treating,
PT preventing and/or prognosing disorders related to the digestive system
PT including digestive system cancers and also for testing and detection
PT e.g. diagnosis -
XX
PS Claim 11; SEQ ID NO 96; 509pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAI99548-AAI99604) and proteins
CC (AAM99936-AAM99984) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 176 AA;

Query Match 76.5%; Score 632; DB 22; Length 176;
Best Local Similarity 94.0%; Pred. No. 8.3e-66;
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFPIVF 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFPIVF 60

Qy 61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
|||||||||| | ||||||||||||||||| ||||||||||||||||||||||||||||||
Db 61 TDALCWIPIFVVKPLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF 120

Qy 121 WHNYRQRKSMDSXG 134
|:|||||||||| |
Db 121 WFNYRQRKSMDSXG 134

RESULT 12
ABS04062
ID ABS04062 standard; Protein; 176 AA.
XX
AC ABS04062;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polypeptide SEQ ID NO 2009.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234697.
PR 25-SEP-2000; 2000US-0234958.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-451937/48.
DR N-PSDB; AAL35644.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis
XX
PS Claim 11; SEQ ID NO 2009; 781pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 176 AA;

```
  Query Match          76.5%;  Score 632;  DB 22;  Length 176;
  Best Local Similarity  94.0%;  Pred. No. 8.3e-66;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AQITYSVAIPLGINLAAPIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKREFPIVP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQITYSVAIPLGINLAAPIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKREFPIVP 60

Qy         61 TDALCWIPIPVVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFKEMIHRF 120
              |||||||||||| | ||||||||||||||||    ||||||||||||||||||||||||||
Db         61 TDALCWIPIPVVKPLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSKG 134
              |:||||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 13
ABB10249
ID   ABB10249 standard; Protein; 176 AA.
XX
AC   ABB10249;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Human cDNA SEQ ID NO: 557.
XX
KW   Human; gene therapy; neural disorder; immune system disorder;
KW   muscular disorder; reproductive disorder; gastrointestinal disorder;
KW   pulmonary disorder; cardiovascular disorder; renal disorder;
KW   proliferative disorder; inflammation.
XX
OS   Homo sapiens.
XX
PN   WO200154474-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01349.
XX
PR   31-JAN-2000; 2000US-179065P.
PR   04-FEB-2000; 2000US-180628P.
PR   24-FEB-2000; 2000US-184664P.
PR   02-MAR-2000; 2000US-186350P.
PR   16-MAR-2000; 2000US-189874P.
PR   17-MAR-2000; 2000US-190076P.
PR   18-APR-2000; 2000US-198123P.
PR   19-MAY-2000; 2000US-205515P.
PR   07-JUN-2000; 2000US-209467P.
PR   28-JUN-2000; 2000US-214886P.
PR   30-JUN-2000; 2000US-215135P.
PR   07-JUL-2000; 2000US-216647P.
PR   07-JUL-2000; 2000US-216880P.
PR   11-JUL-2000; 2000US-217487P.
PR   11-JUL-2000; 2000US-217496P.
PR   14-JUL-2000; 2000US-218290P.
PR   26-JUL-2000; 2000US-220963P.
PR   26-JUL-2000; 2000US-220964P.
PR   14-AUG-2000; 2000US-224518P.
PR   14-AUG-2000; 2000US-224519P.
PR   14-AUG-2000; 2000US-225213P.
PR   14-AUG-2000; 2000US-225214P.
PR   14-AUG-2000; 2000US-225266P.
PR   14-AUG-2000; 2000US-225267P.
PR   14-AUG-2000; 2000US-225268P.
PR   14-AUG-2000; 2000US-225270P.
PR   14-AUG-2000; 2000US-225447P.
PR   14-AUG-2000; 2000US-225757P.
PR   14-AUG-2000; 2000US-225758P.
PR   14-AUG-2000; 2000US-225759P.
PR   18-AUG-2000; 2000US-226279P.
PR   22-AUG-2000; 2000US-226681P.
PR   22-AUG-2000; 2000US-226868P.
PR   22-AUG-2000; 2000US-227182P.
PR   23-AUG-2000; 2000US-227009P.
PR   30-AUG-2000; 2000US-228924P.
PR   01-SEP-2000; 2000US-229287P.
PR   01-SEP-2000; 2000US-229343P.
PR   01-SEP-2000; 2000US-229344P.
PR   01-SEP-2000; 2000US-229345P.
PR   05-SEP-2000; 2000US-229509P.
PR   05-SEP-2000; 2000US-229513P.
PR   06-SEP-2000; 2000US-230437P.
PR   06-SEP-2000; 2000US-230438P.
PR   08-SEP-2000; 2000US-231242P.
PR   08-SEP-2000; 2000US-231243P.
PR   08-SEP-2000; 2000US-231244P.
PR   08-SEP-2000; 2000US-231413P.
PR   08-SEP-2000; 2000US-231414P.
PR   08-SEP-2000; 2000US-232080P.
PR   08-SEP-2000; 2000US-232081P.
PR   12-SEP-2000; 2000US-231968P.
PR   14-SEP-2000; 2000US-232397P.
PR   14-SEP-2000; 2000US-232398P.
PR   14-SEP-2000; 2000US-232399P.
PR   14-SEP-2000; 2000US-232400P.
PR   14-SEP-2000; 2000US-232401P.
PR   14-SEP-2000; 2000US-233063P.
PR   14-SEP-2000; 2000US-233064P.
PR   14-SEP-2000; 2000US-233065P.
PR   21-SEP-2000; 2000US-234223P.
PR   21-SEP-2000; 2000US-234274P.
PR   25-SEP-2000; 2000US-234597P.
PR   25-SEP-2000; 2000US-234598P.
PR   26-SEP-2000; 2000US-235484P.
PR   27-SEP-2000; 2000US-235834P.
PR   27-SEP-2000; 2000US-235836P.
PR   29-SEP-2000; 2000US-236327P.
PR   29-SEP-2000; 2000US-236367P.
PR   29-SEP-2000; 2000US-236368P.
PR   29-SEP-2000; 2000US-236369P.
PR   29-SEP-2000; 2000US-236370P.
PR   02-OCT-2000; 2000US-236802P.
PR   02-OCT-2000; 2000US-237037P.
PR   02-OCT-2000; 2000US-237038P.
PR   02-OCT-2000; 2000US-237039P.
PR   02-OCT-2000; 2000US-237040P.
PR   13-OCT-2000; 2000US-239935P.
PR   13-OCT-2000; 2000US-239937P.
PR   20-OCT-2000; 2000US-240960P.
PR   20-OCT-2000; 2000US-241221P.
PR   20-OCT-2000; 2000US-241785P.
PR   20-OCT-2000; 2000US-241786P.
PR   20-OCT-2000; 2000US-241787P.
PR   20-OCT-2000; 2000US-241808P.
PR   20-OCT-2000; 2000US-241809P.
PR   20-OCT-2000; 2000US-241826P.
PR   01-NOV-2000; 2000US-244617P.
PR   08-NOV-2000; 2000US-246474P.
PR   08-NOV-2000; 2000US-246475P.
PR   08-NOV-2000; 2000US-246476P.
PR   08-NOV-2000; 2000US-246477P.
PR   08-NOV-2000; 2000US-246478P.
PR   08-NOV-2000; 2000US-246523P.
PR   08-NOV-2000; 2000US-246524P.
PR   08-NOV-2000; 2000US-246525P.
PR   08-NOV-2000; 2000US-246526P.
PR   08-NOV-2000; 2000US-246527P.
PR   08-NOV-2000; 2000US-246528P.
PR   08-NOV-2000; 2000US-246532P.
PR   08-NOV-2000; 2000US-246609P.
PR   08-NOV-2000; 2000US-246610P.
PR   08-NOV-2000; 2000US-246611P.
PR   08-NOV-2000; 2000US-246613P.
PR   17-NOV-2000; 2000US-249207P.
PR   17-NOV-2000; 2000US-249208P.
```

```
PR   17-NOV-2000; 2000US-249209P.
PR   17-NOV-2000; 2000US-249210P.
PR   17-NOV-2000; 2000US-249211P.
PR   17-NOV-2000; 2000US-249212P.
PR   17-NOV-2000; 2000US-249213P.
PR   17-NOV-2000; 2000US-249214P.
PR   17-NOV-2000; 2000US-249215P.
PR   17-NOV-2000; 2000US-249216P.
PR   17-NOV-2000; 2000US-249217P.
PR   17-NOV-2000; 2000US-249218P.
PR   17-NOV-2000; 2000US-249244P.
PR   17-NOV-2000; 2000US-249245P.
PR   17-NOV-2000; 2000US-249264P.
PR   17-NOV-2000; 2000US-249265P.
PR   17-NOV-2000; 2000US-249297P.
PR   17-NOV-2000; 2000US-249299P.
PR   17-NOV-2000; 2000US-249300P.
PR   01-DEC-2000; 2000US-250160P.
PR   01-DEC-2000; 2000US-250391P.
PR   05-DEC-2000; 2000US-251030P.
PR   05-DEC-2000; 2000US-251988P.
PR   05-DEC-2000; 2000US-256719P.
PR   06-DEC-2000; 2000US-251479P.
PR   08-DEC-2000; 2000US-251856P.
PR   08-DEC-2000; 2000US-251868P.
PR   08-DEC-2000; 2000US-251869P.
PR   08-DEC-2000; 2000US-251989P.
PR   08-DEC-2000; 2000US-251990P.
PR   11-DEC-2000; 2000US-254097P.
PR   05-JAN-2001; 2001US-259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-476161/51.
DR   N-PSDB; ABA06471.
XX
PT   Isolated nucleic acid molecule encoding an inflammation-associated
PT   polypeptide is used in preventing, treating or ameliorating a medical
PT   condition
XX
PS   Claim 11; SEQ ID NO: 557; 859pp + Sequence Listing; English.
XX
CC   The present invention provides human cDNAs, proteins and related genomic
CC   DNAs. These can be used in the treatment of neural, immune system,
CC   muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC   renal and proliferative disorders and inflammation. The present sequence
CC   is a protein of the invention.
XX
SQ   Sequence  176 AA;

  Query Match           76.5%;  Score 632;  DB 22;  Length 176;
  Best Local Similarity   94.0%;  Pred. No. 8.3e-66;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AQITSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFPFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQITSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFPFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLLTRPFKEMIHRF 120

Qy        121 WHNTRQRKSMDSKG 134
              |:||||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 14
AAU18108
ID   AAU18108 standard; Protein; 176 AA.
XX

AC   AAU18108;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Novel human uterine motility-association polypeptide #15.
XX
KW   Human; uterine motility-association disorder; uterus; pregnancy;
KW   labour; menstrual cycle; gene therapy.
XX
OS   Homo sapiens.
XX
FN   WO200155201-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01317.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230137.
PR   06-SEP-2000; 2000US-0230138.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
```

PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251855.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-488777/53.
DR N-PSDB; AAS28950.
XX
PT Isolated polypeptide and nucleic acid molecules for treating,
PT preventing and/or prognosing disorders related to uterine motility
PT e.g. disorders associated with pregnancy and the menstrual cycle -
XX
PS Claim 11; SEQ ID No 84; 524pp; English.
XX
CC The present invention relates to the isolation of novel human
CC uterine motility-association polypeptides, and cDNA (AAS28936-AAS28994)
CC and genomic sequences encoding for these polypeptides.
CC The sequences of the invention are useful in the diagnosis,
CC treatment, prevention and/or prognosis of diseases associated
CC with uterine motility such as pregnancy and labour, and menstrual
CC disorders. The polynucleotide sequences of the invention are also
CC useful in gene therapy. AAU18094-AAU18152 represent novel human
CC uterine motility-association polypeptides.
CC Note: The reference data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 176 AA;

Query Match 76.5%; Score 632; DB 22; Length 176;
Best Local Similarity 94.0%; Pred. No. 8.3e-66;
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 AQIYSVAIPLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIPLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFKEMIHRF 120
              | |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSKG 134
              |:||||||||||| |
Db        121 WYNYRQRKSMDSKG 134
```

RESULT 15
AAU18344
ID AAU18344 standard; Protein; 176 AA.
XX
AC AAU18344;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human endocrine polypeptide SEQ ID No 299.
XX
KW Endocrine protein; human; mouse; rabbit; goat; horse; food additive;
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;

KW tissue regeneration; anti-infertility.
XX
OS Homo sapiens.
XX
PN WO200155364-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01308.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180626.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 04-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-451936/48.
DR N-PSDB; AAS29573.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders of the endocrine system such as reproductive disorders,
PT endocrine cancers and also for testing and detection e.g. diagnosis -
XX
PS Claim 11; SEQ ID No 299; 604pp; English.
XX
CC Sequences AAU18282-AAU18507 represent endocrine polypeptides of the
CC invention. Endocrine polypeptides and their associated polynucleotides
CC are useful in the diagnosis, treatment and prevention of various types of
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. A pathological condition can be determined by
CC determining the presence or absence of a mutation in an endocrine
CC polynucleotide. The treatable disorders include autoimmune diseases such
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,
CC cerebrovascular disorders such as cerebral ischaemia, nervous system
CC disorders such as Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi, ocular disorders such as corneal infection, endocrine
CC disorders such as premature labour and infertility, gastrointestinal
CC disorders such as Crohn's disease, renal disorders such as
CC glomerulonephritis and respiratory disorders such as asthma. The
CC polypeptides can also be used to aid wound healing, to prevent skin aging
CC due to sunburn, to maintain organs before transplantation, to regenerate
CC tissues and in chemotaxis. The polypeptides can also be used as a food
CC additive or preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

Query Match 76.5%; Score 632; DB 22; Length 176;
Best Local Similarity 94.0%; Pred. No. 8.3e-66;
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 AQIYSVAIPLGINLAAFIITVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVP 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIPLGINLAAFIITVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVP 60

Qy         61 TDALCWIPIPVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||||
Db         61 TDALCWIPIPVVKFLSLLQVEIPGTITSWVVIFILQPINSALNPILYTLTTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSKG 134
              |:|||||||||| |
Db        121 WYNYRQRRSMDSKG 134
```

Search completed: September 25, 2003, 14:59:36
Job time : 42 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 25, 2003, 13:06:05 ; Search time 1645 Seconds
(without alignments)
7003.235 Million cell updates/sec

Title: US-10-049-568-1
Perfect score: 474
Sequence: 1 gcccagatttattcagtggc.....ggagatgccacctgagttaa 474

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	320.4	67.6	841	13	BQ228832	BQ228832 AGENCOURT
2	284	59.9	873	13	BU274328	BU274328 603532573
3	236	49.8	686	13	BU284397	BU284397 603866563
4	199.8	42.2	489	13	BX089223	BX089223 BX089223
5	199.8	42.2	799	29	BZ605862	BZ605862 WHACX477F
6	160.6	33.9	778	13	BQ279098	BQ279098 603865023
7	146	30.8	648	14	CB216867	CB216867 NISC_mq09
c 8	136.4	28.8	744	10	BG304121	BG304121 fl30f05.x
9	131.4	27.7	758	13	BX084029	BX084029 BX084029
10	130	27.4	631	9	AL955885	AL955885 AL955885
11	98.4	20.8	235	9	AW436170	AW436170 75600 WAR
12	90	19.0	198	14	CA997681	CA997681 EST LGR8
13	87.6	18.5	197	10	BP942735	BP942735 EST-CD15N
14	87.6	18.5	197	12	BM280297	BM280297 3'EST-NCD
c 15	71.8	15.1	1078	29	CC244249	CC244249 CH261-120
c 16	69.4	14.6	286	9	AI375172	AI375172 tc10b12.x
17	63.2	13.3	835	28	BE184017	BE184017 CS230-52A
18	60.8	12.8	589	13	BU245340	BU245340 603591301
19	50.6	10.7	456	10	BF021857	BF021857 uy58h08.y
20	49.6	10.5	574	10	BB632910	BB632910 BB632910
21	49.6	10.5	1307	11	AK039086	AK039086 Mus muscu
22	44	9.3	450	14	H17265	H17265 ym37b10.r1
23	44	9.3	511	2	HSM055693	Bx501976 Homo sapi
c 24	43.4	9.2	648	13	BW089852	BW089852 BW089852
25	43.4	9.2	660	13	BW239928	BW239928 BW239928
26	43.4	9.2	680	28	BZ022413	BZ022413 oe138f07.
c 27	43.4	9.2	726	13	BW090495	BW090495 BW090495
c 28	43.4	9.2	733	13	BW086693	BW086693 BW086693
c 29	43.4	9.2	747	13	BW090486	BW090486 BW090486
c 30	43	9.1	602	29	AG228879	AG228879 Lotus jap
31	42.6	9.0	572	28	AQ697633	AQ697633 HS_5525_B
c 32	42.4	8.9	1101	29	CNS016TX	AL107199 Drosophil
33	42.2	8.9	896	29	CNS07DBB	AI440253 T7 end of
34	42.2	8.9	1201	13	BX446356	BX446356 BX446356
35	41.8	8.8	574	29	BZ387557	BZ387557 EINDR49TR
c 36	41.8	8.8	696	13	BW160624	BW160624 BW160624
37	41.8	8.8	739	29	BZ390866	BZ390866 EINDA70TR
38	41.4	8.7	1101	29	CNS0145U	AL103740 Drosophil
c 39	41.2	8.7	779	29	BX229966	BX229966 Danio rer
40	41	8.6	1101	29	CNS0039G	AL063921 Drosophil
c 41	40.6	8.6	208	14	CD429485	CD429485 RTUL_4_D0
42	40.6	8.6	1042	13	BX375647	BX375647 BX375647
43	40.2	8.5	712	13	BW043475	BW043475 BW043475
c 44	40.2	8.5	789	13	BW015323	BW015323 BW015323
45	40	8.4	553	28	AQ775843	HS_2151_B

ALIGNMENTS

RESULT 1
BQ228832
LOCUS BQ228832 841 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7522585 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6051251 5', mRNA sequence.
ACCESSION BQ228832
VERSION BQ228832.1 GI:20410232
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 841)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/ICTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13304 row: m column: 12
High quality sequence start: 87

```
                High quality sequence stop: 603.
FEATURES             Location/Qualifiers
     source          1..841
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6051251"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.  Library constructed by Life
                     Technologies."
BASE COUNT      240 a    160 c    170 g    269 t      2 others
ORIGIN

  Query Match            67.6%;  Score 320.4;  DB 13;  Length 841;
  Best Local Similarity  95.4%;  Pred. No. 7.9e-62;
  Matches  330;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 364

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 424

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTT 484

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
              |||||||||||||||||||||||||||||||||| ||||  |||||||||||||||||||
Db        485 ACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTA 544

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
              ||||||||||||||||||||||||||||||||||||  |  |    ||||||||||||||
Db        545 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCT 604

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAA 346
              ||||||||||||||||||||||||| | ||||||||||| ||  ||
Db        605 TTGAACCCAATTCTCTATACTCTGAACCCAAGACCATTTTAAAGAA 650


RESULT 2
BU274328
LOCUS       BU274328                 873 bp    mRNA    linear   EST 26-NOV-2002
DEFINITION  603532573F1 CSEQCHN53 Gallus gallus cDNA clone ChEST489a19 5', mRNA
            sequence.
ACCESSION   BU274328
VERSION     BU274328.1  GI:25545278
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 873)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology (UMIST
            )
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1..873
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Compton Line 15I"
                     /db_xref="taxon:9031"
                     /clone="ChEST489a19"
                     /sex="Female"
                     /tissue_type="cerebrum"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN53"
                     /note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
                     EcoRI; Site_2: NotI; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with EcoRI
                     , size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
BASE COUNT      252 a    190 c    160 g    271 t
ORIGIN

  Query Match            59.9%;  Score 284;  DB 13;  Length 873;
  Best Local Similarity  79.2%;  Pred. No. 1.2e-53;
  Matches  374;  Conservative    0;  Mismatches   95;  Indels    3;  Gaps    3;

Qy          4 CAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTT 63
              ||||||||||| ||| | |||||| | ||| | || ||||| || ||| ||||||| || 
Db         42 CAGATTTATTCTGTGTC-ATTTTTTTAGGTGTAAACTTGGCAGCTTTTCTCATCAT-GTC 99

Qy         64 TTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATA 123
              |||||||||||||| ||||| || |||||||| ||||  || || |  || |||||||| 
Db        100 TTTTCCTATGGAAGTATGTTCTACAGTGTTCACCAAACAGCTATTATGGCCACTGAAATT 159

Qy        124 CGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTACT 183
              | ||||||  ||||||||||||||||||  |||||||||||||  ||||||||| || ||||||||
Db        160 CAGAATCATATTAAAAAAGAGATGATACTTGCCAAACGCTTTTTCTTCATTGTATTTACC 219

Qy        184 GATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAA 243
              |||||| | || ||||||||||||||||| |    |||||  |||||   | || || ||||||
Db        220 GATGCACTTTGTTGGATACCCATTTTTATTTTGAAACTCCTTTCTTTACTGCAAGTAGAA 279

Qy        244 ATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTG 303
              ||||||||||||||||| |||||||||||||||   |    |    | || || |||||| ||
Db        280 ATACCAGGTACCATAACATCTTGGGTAGTGATTTTTATCTTGCCAATAAATAGTGCTCTG 339

Qy        304 AACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGG 363
              || ||  ||||||| ||| |||| ||| |||| || |||||||||||||| | |||||||
Db        340 AATCCTCTTCTCTACACTTTGACTACACGACCCTTCAAAGAAATGATTCACCAGTTTTGG 399

Qy        364 CATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTA-TCAGAAAACATATGCTCCA 422
               | |||||||||||||||| ||| | |    |||||| | |||||||| | |||| ||||
Db        400 TACAACTACAGACAAAGAAGATCCAAGAGAGGCAAAGGCAGTCAGAAAGCTTATGGTCCA 459

Qy        423 TCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              || |||||| |||| || |||||||||| |||| || ||  |||| || |||
Db        460 TCCTTCATTTGGGTAGAGATGTGGCCAATGCATGAAATCACACCAAAGCTAA 511


RESULT 3
BU284397
LOCUS       BU284397                 686 bp    mRNA    linear   EST 27-NOV-2002
DEFINITION  603866563F1 CSEQCHN54 Gallus gallus cDNA clone ChEST886m16 5', mRNA
            sequence.
ACCESSION   BU284397
```

```
VERSION     BU284397.1  GI:25733853
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 686)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology (UMIST
            )
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1..686
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Compton Line 15I"
                     /db_xref="taxon:9031"
                     /clone="ChEST886m16"
                     /sex="Female"
                     /tissue_type="not cerebrum or cerebellum"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN54"
                     /note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
                     EcoRI; Site_2: NotI; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with EcoRI
                     , size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer
                     reannealing hybridization was used."
BASE COUNT      207 a    151 c    131 g    196 t      1 others
ORIGIN

  Query Match             49.8%;  Score 236;  DB 13;  Length 686;
  Best Local Similarity   78.4%;  Pred. No. 7.8e-43;
  Matches  345;  Conservative    0;  Mismatches   90;  Indels    5;  Gaps    5;

Qy         36 TAATTTGGCCGCATTTATCATCATAGTTTTTCCTATGGAAGCATGTTTTATAGTGTTCA 95
              |||  |||| || | | ||||||| || |||||||||||||| ||||| || ||||||||
Db          2 TAAACTGGCAGCTTGTCTCATCAT-GTCTTTTCCTATGGAAGTATGTTCTACAGTGTTCA 60

Qy         96 TCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGC 155
                |||| || || |  || |||||||| | ||||||  ||||||||||||||| | |||
Db         61 CCAAACAGCTATTATGGCCACTGAAATTCAGAATCATATTAAAAAAGAGATGATAC-TGC 119

Qy        156 CAAACGTTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTTGTAGC 215
              |||||| |||||||| || |||||||| ||||| | || |||||||||||||| |
Db        120 CAAACGCTTTTTCTTCATTGTATTTACCGATGCACTTTG-TGGATACCCATTTTTATTTT 178

Qy        216 GAAACCTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTAGTGAT 275
              |||||  ||||| | || || ||||||||||||||||||||||| || |||||||||||
Db        179 GAAACTCCTTTCTTTACTGCAAGTAGAAATACCAGGTACCATAACATC-TGGGTAGTGAT 237

Qy        276 TGGTTATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACC 335
              | |  |    | || || |||||| |||| ||  ||||||| ||| |||| ||| ||||

Db        238 TTTTATCTTGCCAATAAATAGTGCTCTGAATCCTCTTCTCTACACTTTGACTACACGACC 297

Qy        336 ATTTAAAGAAATGATTCATCGGTTTTGGCATAACTACAGACAAAGAAAATCTATGGACAG 395
               || ||||||||||||| | |||||| | ||||||||||||||| ||| | |      |
Db        298 CTTCAAAGAAATGATTCACCAGTTTTGGTACAACTACAGACAAAGAAGATCCAAGAGAGG 357

Qy        396 CAAAGGTA-TCAGAAAACATATGCTCCATCATTCATCTGGGGGGAAATGTGGCCACTGCA 454
              |||||| | ||||||| | |||| |||||| ||||| ||||    || ||||||||| ||||
Db        358 CAAAGGCAGTCAGAAAGCTTATGGTCCATCCTTCATTTGGGTAGAGATGTGGCCAATGCA 417

Qy        455 GGAGATGCCACCTGAGTTAA 474
              || ||  ||||  || |||
Db        418 TGAAATCACACCAAAGCTAA 437


RESULT 4
BX089223
LOCUS       BX089223                 489 bp    mRNA    linear   EST 23-JAN-2003
DEFINITION  BX089223 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGp998M245082 ;
            IMAGE:2063423, mRNA sequence.
ACCESSION   BX089223
VERSION     BX089223.1  GI:27825935
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998M245082.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1..489
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998M245082 ; IMAGE:2063423"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      166 a     95 c     86 g    142 t
ORIGIN

  Query Match             42.2%;  Score 199.8;  DB 13;  Length 489;
```

```
  Best Local Similarity   88.7%;  Pred. No. 1.1e-34;
  Matches  228;  Conservative    0;  Mismatches   27;  Indels    2;  Gaps    1;

Qy       218 AACCTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTG 277
             || | ||   ||| ||||  |   |  |||||||||||||||||||||||||||||||||
Db       147 AAACACTAAACTATTTCATTTTTCTTTTTAGGTACCATAACCTCTTGGGTAGTGATTT 206

Qy       278 GTTATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCAT 337
             |  |  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       207 TTATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCAT 266

Qy       338 TTAAAGAAATGATTCATCGGTTTTGGCATAACTACAGACAAAGAAAATCTATGGACAGCA 397
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       267 TTAAAGAAATGATTCATCGGTTTTGGTATAACTACAGACAAAGAAAATCTATGGACAGCA 326

Qy       398 AAGGTATCAGAAAACATATGCTCCATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGA 457
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       327 AAGG--TCAGAAAACATATGCTCCATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGA 384

Qy       458 GATGCCACCTGAGTTAA 474
             ||||||||||||||||
Db       385 GATGCCACCTGAGTTAA 401



RESULT 5
BZ605862
LOCUS       BZ605862               799 bp    DNA     linear   GSS 08-JUN-2003
DEFINITION  WHACX477F Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-18H22, genomic survey sequence.
ACCESSION   BZ605862
VERSION     BZ605862.1  GI:31514423
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 799)
  AUTHORS   Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
  TITLE     ESP: a sequence-based approach to the structural genomics of tumors
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Volik SV
            Colin Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..799
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="MCF7_1-18H22"
                     /sex="female"
                     /clone_lib="Human MCF7 breast cancer cell line library
                     (MCF7_1)"
                     /note="Vector: pECBAC1; Site_1: HindIII; This library was
                     constructed from MCF7 breast cancer cell line by Amplicon
                     Express (http://www.genomex.com) using their standard
                     procedure."
BASE COUNT      277 a    145 c    127 g    250 t
ORIGIN

  Query Match           42.2%;  Score 199.8;  DB 29;  Length 799;
  Best Local Similarity   88.7%;  Pred. No. 1.1e-34;
  Matches  228;  Conservative    0;  Mismatches   27;  Indels    2;  Gaps    1;

Qy       218 AACCTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTG 277
             || | ||   ||| ||||  |   |  |||||||||||||||||||||||||||||||||
Db       389 AAACACTAAACTATTTCATTTTTCTTTTTAGGTACCATAACCTCTTGGGTAGTGATTT 448

Qy       278 GTTATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCAT 337
             |  |  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       449 TTATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCAT 508

Qy       338 TTAAAGAAATGATTCATCGGTTTTGGCATAACTACAGACAAAGAAAATCTATGGACAGCA 397
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       509 TTAAAGAAATGATTCATCGGTTTTGGTATAACTACAGACAAAGAAAATCTATGGACAGCA 568

Qy       398 AAGGTATCAGAAAACATATGCTCCATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGA 457
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       569 AAGG--TCAGAAAACATATGCTCCATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGA 626

Qy       458 GATGCCACCTGAGTTAA 474
             ||||||||||||||||
Db       627 GATGCCACCTGAGTTAA 643



RESULT 6
HU279098
LOCUS       HU279098               778 bp    mRNA    linear   EST 27-NOV-2002
DEFINITION  603865023F1 CSEQCHN54 Gallus gallus cDNA clone ChEST882n6 5', mRNA
            sequence.
ACCESSION   HU279098
VERSION     HU279098.1  GI:25728552
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 778)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology (UMIST
            )
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360408
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1..778
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Compton Line 15I"
                     /db_xref="taxon:9031"
                     /clone="ChEST882n6"
                     /sex="female"
                     /tissue_type="not cerebrum or cerebellum"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQCHN54"
                     /note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
                     EcoRI; Site_2: NotI; This normalized library was
                     constructed from 1 million independent clones. cDNA
                     synthesis was initiated using an oligo(dT) primer, using
                     methylated C in the first strand synthesis reaction.
                     Following this first strand reaction, double-stranded cDNA
                     was blunted, ligated to NotI adapters, digested with EcoRI
                     , size-selected, and cloned into the NotI and EcoRI
                     compatible sites of a custom modified MCS of the
                     pBluescript (KS+) vector. The library was normalized in 2
                     rounds using conditions adapted from Soares et al., PNAS
                     (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
                     (1996): 791, except that a significantly longer

  Query Match           42.2%;  Score 199.8;  DB 29;  Length 778;
  Best Local Similarity   88.7%;  Pred. No. 1.1e-34;
  Matches  228;  Conservative    0;  Mismatches   27;  Indels    2;  Gaps    1;

Qy       218 AACCTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTG 277
             || | ||   ||| ||||  |   |  |||||||||||||||||||||||||||||||||
Db       389 AAACACTAAACTATTTCATTTTTCTTTTTAGGTACCATAACCTCTTGGGTAGTGATTT 448

Qy       278 GTTATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCAT 337
             |  |  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       449 TTATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCAT 508

Qy       338 TTAAAGAAATGATTCATCGGTTTTGGCATAACTACAGACAAAGAAAATCTATGGACAGCA 397
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       509 TTAAAGAAATGATTCATCGGTTTTGGTATAACTACAGACAAAGAAAATCTATGGACAGCA 568

Qy       398 AAGGTATCAGAAAACATATGCTCCATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGA 457
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       569 AAGG--TCAGAAAACATATGCTCCATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGA 626

Qy       458 GATGCCACCTGAGTTAA 474
             ||||||||||||||||
Db       627 GATGCCACCTGAGTTAA 643
```

```
                    reannealing hybridization was used."
BASE COUNT      247 a    167 c    141 g    223 t
ORIGIN

  Query Match             33.9%;  Score 160.6;  DB 13;  Length 778;
  Best Local Similarity   77.6%;  Pred. No. 7.2e-26;
  Matches  232;  Conservative    0;  Mismatches   64;  Indels    3;  Gaps    3;

Qy        177 ATTTACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCA 236
              ||||| ||||| | || ||||||||||||||| |   |||||  |||||  | || ||
Db          1 ATTTACCGATGCACT-TGTTGGATACCCATTTTTATTTTGAAACTCCTTTCTTTACTGCA 59

Qy        237 GGTAGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAG 296
               ||||||||||||||||||||| ||||||||||||||  |    |    | || || ||
Db         60 AGTAGAAATACCAGGTACCATAACATCTTGGGTAGTGATTTTTATCTTGCCAATAAATAG 119

Qy        297 TGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCG 356
              |||| |||| ||   ||||||| ||| |||| ||| |||| || |||||||||||| |
Db        120 TGCTCTGAATCCTCTTCTCTACACTTTGACTACACGACCCTTCAAAGAAATGATTCACCA 179

Qy        357 GTTTTGGCATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTA-TCAGAAAACATA 415
              | ||||| | ||||||||||||||| ||| | |      ||||||| | ||||||| | ||
Db        180 G-TTTGGTACAACTACAGACAAAGAAGATCCAAGAGAGGCAAAGGCAGTCAGAAAGCTTA 238

Qy        416 TGCTCCATCATTCATCTGGGGGGAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAA 474
              || |||||| ||||| ||||  || ||||||||| |||| || ||  ||||  || |||
Db        239 TGGTCCATCCTTCATTTGGGTAGAGATGTGGCCAATGCATGAAATCACACCAAAGCTAA 297
```

```
RESULT 7
CB216867
LOCUS       CB216867                 648 bp    mRNA    linear   EST 06-FEB-2003
DEFINITION  NISC_nq09g08.y1 NICHD_HS_Ut2 Homo sapiens cDNA clone IMAGE:5938742
            5', mRNA sequence.
ACCESSION   CB216867
VERSION     CB216867.1  GI:28265059
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation:
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM13167  row: M  column: 15
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..648
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5938742"
                     /sex="female"
                     /tissue_type="normal endometrium, mid-secretory phase,
                     cycle day 23"
                     /lab_host="DH10B (T1-resistant)"
                     /clone_lib="NICHD_HS_Ut2"
                     /note="Organ: uterus; Vector: pCMV-SPORT6.1.ccdb (ResGen,
                     Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV;
                     Cloned unidirectionally from microquantity amounts of mRNA
                     from normal endometrial tissue (mid-secretory phase, cycle
                     day 23). Average insert size 1.6 kb. Library constructed
                     by ResGen (Invitrogen Corporation)."
BASE COUNT      180 a    112 c    133 g    223 t
ORIGIN

  Query Match             30.8%;  Score 146;  DB 14;  Length 648;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-22;
  Matches  146;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 562

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 622

Qy        121 ATACGGAATCAAGTTAAAAAAGAGAT 146
              ||||||||||||||||||||||||||
Db        623 ATACGGAATCAAGTTAAAAAAGAGAT 648
```

```
RESULT 8
BG304121/c
LOCUS       BG304121                 744 bp    mRNA    linear   EST 23-FEB-2001
DEFINITION  fl30f05.x1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone
            IMAGE:3815552 3' similar to TR:Q9VBP0 Q9VBP0 CG5042 PROTEIN. ;,
            mRNA sequence.
ACCESSION   BG304121
VERSION     BG304121.1  GI:13101648
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Sugano,S., Kawakami,K., Johnson,S., Li,F., Marra,M., Eddy,S.,
            Hillier,L., Clifton,S., Allen,M., Gibbons,M., Jost,S., Kucaba,T.,
            Martin,J., Pape,D., Steptoe,M., Underwood,K., Theising,B., Ritter
            ,E., Bowers,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1999
  JOURNAL   Unpublished
COMMENT     Contact: S.L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution information can be found through the I.M.A.G.E.
            Consortium/LLNL, send email to: info@image.llnl.gov
            Seq primer: T7 from Gibco
            High quality sequence stop: 510.
FEATURES             Location/Qualifiers
     source          1..744
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /strain="AB"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:3815552"
                     /sex="mixed (one male and one female, including
                     unfertilized eggs)"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage resistant)"
                     /clone_lib="Sugano Kawakami zebrafish DRA"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);
                     Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer  [ATGTGGCCTTTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a  DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
```

insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science) and kindly donated by Dr. Koichi Kawakami. Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA. "

BASE COUNT 204 a 158 c 186 g 196 t
ORIGIN

Query Match 28.8%; Score 136.4; DB 10; Length 744;
Best Local Similarity 62.1%; Pred. No. 2e-20;
Matches 215; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

Qy 25 TTTCTTGGTATTAATTGGCCGCATTTATCATCATAGTTTTTCCTATGGAAGCATGTTT 84
Db 744 TTCATGGGACTTACCCGGCTGCTTTCCTTCATCATCGTGGTCTCGTACTCCAGCATGTTT 685

Qy 85 TATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAG 144
Db 684 TACTCCATCTATAAGACGGGCATCAGAGCCACAGAGCTGCCGGGTCGACTGCACAGAGAC 625

Qy 145 ATGATCCTTGGCAAACGTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCC 204
Db 624 GTGGCCATGGCACACCGCTTCTTCTTCATCGTGTTTTCTGATGCCTTCTGCTGGATTCCC 565

Qy 205 ATTTTTGTAGCGAAACCTCTTTCTCACTGCTTCAGCTAGAAATACCAGGTACCATAACCTCT 264
Db 564 ATATTCATGGTCAAAATCCTCTCTCTAATGGAGGTGGAGATACCAGGAACCATCACATCT 505

Qy 265 TGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTTATACTCTG 324
Db 504 TGGGTGGTGATCTTCATCCTGCCCATCAACAGTGCCTTAAACCCCATCCTTTACACACTG 445

Qy 325 ACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGCATAACT 370
Db 444 ACCACCAGTTTTTTCAGAGACGAGTGGAACATTTATGGGGACACT 399

RESULT 9
BX084029
LOCUS BX084029 758 bp mRNA linear EST 25-APR-2003
DEFINITION BX084029 AGENAE Rainbow trout normalized multi-tissues library (tcac) Oncorhynchus mykiss cDNA clone tcac0004c.h.01 5prim, mRNA sequence.
ACCESSION BX084029
VERSION BX084029.1 GI:27747446
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 758)
AUTHORS Govoroun,M., Guiguen,Y. and Le Gac,F.
TITLE Construction and primary characterization of normalized cDNA libraries in rainbow trout, Oncorhynchus mykiss
JOURNAL Unpublished
COMMENT Contact: Guiguen Y
INRA - SCRIBE
Campus de beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptator and repetitions. Contact us at eigenasupport@jouy.inra.fr to obtain the chromatogram of this sequence.
Plate: 0004 row: h column: 1
Seq primer: T7.
FEATURES Location/Qualifiers
source 1. .758
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="tcac0004c.h.01"
/tissue_type="adipose tissue, blood, brain, differentiating gonads, gills, interrenal, intestine, kidney, liver, muscle, ovary, pituitary, testis"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Rainbow trout normalized multi-tissues library (tcac)"
/note="Vector: pT723D-pac; Clone distribution : AGENAE Resource centre. Francois PIUMI, Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, France"

BASE COUNT 165 a 217 c 178 g 198 t
ORIGIN

Query Match 27.7%; Score 131.4; DB 13; Length 758;
Best Local Similarity 61.6%; Pred. No. 2.7e-19;
Matches 210; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

Qy 10 TAGTCAGTGGCAATTTTTCTTGGTATTAATTTGGCGCATTTATCATCATAGTTTTTCC 69
Db 282 TACTCTATCGGCATATTCCTGGGTCGAACCTGGCAGCATTCCTAGTGATCGTGTTCTCT 341

Qy 70 TATGGAAGCATGTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAAT 129
Db 342 TACTCCTCCATGTTCTACTCCATCTACAAGACTGGTATCAACGCTACAGACCTGAGGAGC 401

Qy 130 CAAGTTAAAAAAGAGATGATCCTTGGCAAACGTTTTTCTTTATAGTATTTACTGATGCA 189
Db 402 AGGCTACACAGAGACCTGGCTGTGGCTAACAGGTTCTTCTTCATCGTCTTCTCTGACGCC 461

Qy 190 TTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCA 249
Db 462 CTCTGCTGGATCCCCATCTTCCTGGTCAAGATACTGTCCCTGCTGCAGGTCGAGATACCA 521

Qy 250 GGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCTTTGAACCCA 309
Db 522 GGTACCATCTCATCCTGGGTGGTGATCTTTATATTACCCATCAACAGTGCCCTGAACCCC 581

Qy 310 ATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGAT 350
Db 582 ATCCTCTACACCCTGACTACCAGCTTCTCAGAGAACAGGT 622

RESULT 10
AL955885
LOCUS AL955885 631 bp mRNA linear EST 27-NOV-2002
DEFINITION AL955885 XGC-gastrula Silurana tropicalis cDNA clone TGas123k02 5', mRNA sequence.
ACCESSION AL955885
VERSION AL955885.1 GI:25779480
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 631)
AUTHORS Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2002
JOURNAL Unpublished
COMMENT Contact: Taylor R
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TGasl23k02.plcSp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.
FEATURES Location/Qualifiers
source 1. .631

```
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1232
;   LENGTH: 1557
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (16)
;   OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-134-001C-1232

  Query Match             8.0%;  Score 37.8;  DB 4;  Length 1557;
  Best Local Similarity   54.7%;  Pred. No. 0.5;
  Matches   75;  Conservative    0;  Mismatches   62;  Indels    0;  Gaps    0;

Qy        337 TTTAAAGAAATGATTCATCGGTTTTGGCATAACTACAGACAAAGAAAATCTATGGACAGC 396
              || || ||    ||  ||  || ||| ||| | ||||| | | | | |||| ||||
Db       1293 TTCAACGAGTAATTTACTCTTTTGTGGGATACCACCAGACGCATACACTGTATGAACAGG 1234

Qy        397 AAAGGTATCAGAAAACATATGCTCCATCATTCATCTGGGGGCAAATGTGGCCACTGCAGG 456
              ||    || |       |  || | ||| ||| || | |    ||||  |   || ||
Db       1233 AATATGATTATCTTCAAATTGTTTCATAATTAATTTTGTTCCAAATGCTGTAGCTTCAAT 1174

Qy        457 AGATGCCACCTGAGTTA 473
              | ||||   ||| | |
Db       1173 ATATGCTCGATGAATCA 1157


Search completed: September 25, 2003, 13:59:13
Job time : 59 secs
```

```
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="pLGR8"
                     /sex="female"
                     /tissue_type="uterus"
                     /dev_stage="neonatal day 14"
                     /clone_lib="Sus scrofa neonatal uterus RT-PCR"
                     /note="Organ: uterus; cDNA generated by RT-PCR from total
                     RNA extracted from porcine neonatal uterine tissue."
BASE COUNT       42 a     60 c     36 g     60 t
ORIGIN

  Query Match             19.0%;  Score 90;  DB 14;  Length 198;
  Best Local Similarity   66.5%;  Pred. No. 5.8e-10;
  Matches  129;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy        178 TTTACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAG 237
                || |||||||  | ||||||||| ||| | |||||      |||  ||| || ||  | | |
Db          2 TTCTCTGATGCCATCTGCTGGATTCCCGTATTTGTCATTAAAATTCTCTCCCTCTTCCGG 61

Qy        238 GTAGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGT 297
              || |||||||||||  |  | || || ||||| || |||  ||       | |  | ||||||
Db         62 GTGGAAATACCAGGCTCACTCACTTCCTGGGTGGTCATTTTTTTCCTTCCGGTGAACAGT 121

Qy        298 GCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGG 357
              || ||||||||||| ||||| ||||| || || ||       || || || | | |  | | |
Db        122 GCCTTGAACCCAATCCTCTACACTCTCACAACCAGCTTTTTCAAGGACACGTTGAAACAG 181

Qy        358 TTTTGGCATAACTA 371
                |   ||| ||| |
Db        182 CTGCTGCACAACCA 195


RESULT 13
BF942735
LOCUS       BF942735                 197 bp    mRNA    linear   EST 15-SEP-2001
DEFINITION  EST-CD15N-040 human CD15+ myeloid progenitor cells cDNA Library
            Homo sapiens cDNA 3', mRNA sequence.
ACCESSION   BF942735
VERSION     BF942735.1  GI:15624134
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 197)
  AUTHORS   Chen,J., Lee,S., Zhou,G., Rowley,J.D. and Wang,S.M.
  TITLE     A high-throughput GLGI procedure for converting large number of
            SAGE tag sequences into 3' ESTs
  JOURNAL   Unpublished
COMMENT     Contact: Wang SM
            Hem/Onc
            University of Chicago Medical Center
            5841 S. Maryland Ave., MC2115, Chicago, IL 60637, USA
            Tel: 773-702-6788
            Fax: 773-702-3002
            Email: swang1@midway.uchicago.edu
            This EST fragment was amplified from human CD15+ myeloid progenitor
            cells cDNA Library with GLGI technique (Generation of Longer cDNA
            fragments from SAGE tags for Gene Identification, Proc. Natl. Acad.
            Sci. USA 97, 349, 2000), which starts from the 3' end till the last
            CATG site of the target cDNA sequence.
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1..197
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="bone marrow"
                     /cell_type="CD15+ myeloid progenitor cells"
                     /clone_lib="human CD15+ myeloid progenitor cells cDNA
                     Library"
BASE COUNT       44 a     38 c     37 g     78 t
ORIGIN

  Query Match             18.5%;  Score 87.6;  DB 10;  Length 197;
  Best Local Similarity   66.3%;  Pred. No. 2e-09;
  Matches  126;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy         62 TTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAA 121
              ||||| |||||    |  |||||| | |         ||||  |||  ||| |       | || |||
Db          8 TTTTTCCCTATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAG 67

Qy        122 TACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTTA 181
              || |||||     ||   || |||| ||      |||| || ||||| || |||||||| ||
Db         68 TAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTATAGTGTTCT 127

Qy        182 CTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAG 241
              |||||||  | |||||||| ||  | |||||||  |||   ||||| ||  | | ||| |
Db        128 CTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGG 187

Qy        242 AAATACCAGG 251
              |||||||| |
Db        188 AAATACCATG 197


RESULT 14
BM280297
LOCUS       BM280297                 197 bp    mRNA    linear   EST 17-SEP-2002
DEFINITION  3'EST-NCD15-045 human CD15+ myeloid progenitor cells cDNA Library
            Homo sapiens cDNA 3', mRNA sequence.
ACCESSION   BM280297
VERSION     BM280297.1  GI:23025830
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 197)
  AUTHORS   Chen,J., Sun,M., Lee,S., Zhou,G., Rowley,J.D. and Wang,S.M.
  TITLE     Identifying novel transcripts and novel genes in the human genome
            by using novel SAGE tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12257-12262 (2002)
COMMENT     Contact: Wang SM
            Hem/Onc
            University of Chicago Medical Center
            5841 S. Maryland Ave., MC2115, Chicago, IL 60637, USA
            Tel: 773-702-6788
            Fax: 773-702-3002
            Email: swang1@midway.uchicago.edu
            This EST fragment was amplified from human CD15+ myeloid progenitor
            cells cDNA Library with GLGI technique (Generation of Longer cDNA
            fragments from SAGE tags for Gene Identification, Proc. Natl. Acad.
            Sci. USA 97, 349, 2000; A high-throughput GLGI procedure for
            converting a large number of SAGE tag sequences into 3' ESTs, Genes
            , Chromosomes & Cancers, In press), which starts from the 3' end
            till the last CATG site of the target cDNA sequence.
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1..197
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="bone marrow"
                     /cell_type="CD15+ myeloid progenitor cells"
                     /clone_lib="human CD15+ myeloid progenitor cells cDNA
                     Library"
BASE COUNT       44 a     38 c     37 g     78 t
ORIGIN

  Query Match             18.5%;  Score 87.6;  DB 12;  Length 197;
  Best Local Similarity   66.3%;  Pred. No. 2e-09;
  Matches  126;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy         62 TTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAA 121
```

```
                    /organism="Silurana tropicalis"
                    /mol_type="mRNA"
                    /db_xref="taxon:8364"
                    /clone="TGas123k02"
                    /dev_stage="gastrula (stages 10.5-13 mixed)"
                    /lab_host="Escherichia coli XL1-blue"
                    /clone_lib="XGC-gastrula"
                    /note="Vector: pCS107: Site_1: EcoRI; Site_2: NotI: cDNA
                    was oligo dT primed from 5ug of poly A+ RNA from stages
                    10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
                    into pCS107 with EcoRI at the 5' end and NotI at the 3'
                    end."
BASE COUNT      180 a    166 c    130 g    154 t      1 others
ORIGIN

  Query Match             27.4%;  Score 130;  DB 9;  Length 631;
  Best Local Similarity   60.8%;  Pred. No. 5.6e-19;
  Matches  230;  Conservative    0;  Mismatches  145;  Indels    3;  Gaps    1;

Qy          1 GCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATA 60
Db         80 GCCCGGGGATATTCCGTGAGCATATTCCTTGGACTGAACTTACTGGCTTTCGTCACCATC 139

Qy         61 GTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAA 120
Db        140 GTCTTCTCCTACAGCAGCATGTTCTACTCCATTCATAAGACTGGAGCCAAAACAGCAGA- 198

Qy        121 ATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTCTTTATAGTATTT 180
Db        199 --ACGCAGCGTCCTGTCACGAGAGGTCACCATTGCCAAGAGGTTCTTCTTCATAGTTTTC 256

Qy        181 ACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTA 240
Db        257 ACCGATGCTCTCTGCTGGATCCCTATCTTCTTACTAAAAGCAGTTTCTCTCACGGAAGCA 316

Qy        241 GAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGTTATTCTGCCATTAACAGTGCT 300
Db        317 GAGATTCCAGGCACCATCACTTCATGGATTGTGATCTTCATCCTGCCTATTAACAGTGCT 376

Qy        301 TTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTT 360
Db        377 CTGAACCCCATATTGTACACACTCACCACCTCCTCTTTCCAGGAGAAGCTAAAACAGTGT 436

Qy        361 TGGCATAACTACAGACAA 378
Db        437 CTACAGAGCAAAAGGCAA 454


RESULT 11
AW436170
LOCUS       AW436170              235 bp    mRNA    linear   EST 09-JUL-2000
DEFINITION  75600 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   AW436170
VERSION     AW436170.1  GI:6971476
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 235)
  AUTHORS   Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
            Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush
            ,J. and Keele,J.W.
  TITLE     Porcine gene discovery by normalized cDNA-library sequencing and
            EST cluster assembly
  JOURNAL   Mamm. Genome 13 (8), 475-478 (2002)
  MEDLINE   22213789
   PUBMED   12226715
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 20
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 34  row: F  column: 22
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1..235
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 2PIG"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from testis, ovary,
                     endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT       49 a     64 c     45 g     77 t
ORIGIN

  Query Match             20.8%;  Score 98.4;  DB 9;  Length 235;
  Best Local Similarity   66.5%;  Pred. No. 7.5e-12;
  Matches  141;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy        160 CGTTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAA 219
Db          5 CTTTTTTTTTTTATAGGGTTCTCTGATGCCATCTGCTGGATTCCCGTATTTGTCATTAAA 64

Qy        220 CCTCTTTCACTGCTTCAGGTAGAAATACCAGGTACCATAACCTCTTGGGTAGTGATTGGT 279
Db         65 ATTCTCTCCCTCTTCCGGGTGGAAATACCAGGCTCACTCACTTCCTGGGGGGTCATTTTT 124

Qy        280 TATTCTGCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTT 339
Db        125 ITCCTTCCGGTGAACAGTGCCTTGAACCCAATCCTTTACACTCTCACAACCAGCTTTTTC 184

Qy        340 AAAGAAATGATTCATCGGTTTTGGCATAACTA 371
Db        185 AAGGACACGTTGAAACAGCTGCTGCACAACCA 216


RESULT 12
CA997681
LOCUS       CA997681              198 bp    mRNA    linear   EST 08-JAN-2003
DEFINITION  EST LGR8 Sus scrofa neonatal uterus RT-PCR Sus scrofa cDNA clone
            pLGR8 similar to Porcine LGR8 (INSL3 receptor), mRNA sequence.
ACCESSION   CA997681
VERSION     CA997681.1  GI:27545549
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 198)
  AUTHORS   Bartol,F.F., Wiley,A.A., Bathgate,R.A.D. and Bagnell,C.A.
  TITLE     INSL3 receptor (LGR8) expression by neonatal porcine uterine
            tissues
  JOURNAL   Unpublished
COMMENT     Contact: Bartol FF
            Department of Animal Sciences
            Auburn University, Cellular and Molecular Biosciences Program
            240 Upchurch Hall, Auburn, AL 36849-5415, USA
            Tel: 334 844 1506
            Fax: 334 844 1519
            Email: bartoff@auburn.edu
            Sequence of cDNA (198 bp) generated by RT-PCR from total RNA
            extracted from porcine neonatal uterine tissues obtained on
            postnatal day 14.
FEATURES             Location/Qualifiers
     source          1..198
```

```
                 ||||| |||||  |  ||||| | |    ||||  |||  ||| |      | || |||
Db           8 TTTTTCCCTATATTACTATGTTCTGTTCCATTCAAAAAACCGCCTTGCAGACCACAGAAG 67

Qy         122 TACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGTTTTTTCTTTATAGTATTTA 181
                || |||||    ||  || |||| ||     |||| || ||||| || |||||||| ||
Db          68 TAAGGAATTGTTTTGGAAGAGAGGTGGCTGTTGCAAATCGTTTCTTTTTATAGTGTTCT 127

Qy         182 CTGATGCATTATGCTGGATACCCATTTTTGTAGCGAAACCTCTTTCACTGCTTCAGGTAG 241
               ||||||  | |||||||| ||  | |||||||  |||    ||||| ||  | | ||| |
Db         128 CTGATGCCATCTGCTGGATTCCTGTATTTGTAGTTAAAATCCTTTCCCTCTTCCGGGTGG 187

Qy         242 AAATACCAGG 251
               |||||||| |
Db         188 AAATACCATG 197


RESULT 15
CC244249/c
LOCUS       CC244249                1078 bp    DNA     linear   GSS 12-MAY-2003
DEFINITION  CH261-1205_RM1.1 CH261 Gallus gallus genomic clone CH261-1205,
            genomic survey sequence.
ACCESSION   CC244249
VERSION     CC244249.1  GI:30570912
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1078)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 34
            High quality sequence stop: 532.
FEATURES             Location/Qualifiers
     source          1..1078
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-1205"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
BASE COUNT      240 a    176 c    141 g    219 t    302 others
ORIGIN

  Query Match             15.1%;  Score 71.8;  DB 29;  Length 1078;
  Best Local Similarity   63.6%;  Pred. No. 6.8e-06;
  Matches  143;  Conservative    0;  Mismatches   77;  Indels    5;  Gaps    2;

Qy          31 GGTATTAATTTGGCCGCATTTATCATCATAGTTTTTCCTATGGAAGCATGTTTTATAGT 90
               ||| | ||  ||    || || ||||||| ||| ||| | || |   |||||  ||| |
Db         296 GGTGTGAACCTGCATGCCTTCATCATCACAGTGTTTCCATAGGCCAGCATACTTTGTTCC 237

Qy          91 GTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATC 150
                 ||||  ||| || | |     |||    ||| |  |||||||  || | |  ||| | |
Db         236 ATTCAAAAAACTGTCCTTCAGGCA---GAAGTGAGGAATCACATTCACAGGGAGGTTACT 180

Qy         151 CTTGCCAAACGT--TTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTT 208
                 ||||||| |     ||||| ||| ||||||| ||||||||     |||||||| |   | |
Db         179 GTTGCCAATCAATATTTTTTTTTCTAGTATTCACTGATGCCACCTGCTGGATTCGTGTCT 120

Qy         209 TTGTAGCGAAACCTCTTTCACTGCTTCAGGTAGAAATACCAGGTA 253
               ||||     ||| | |||||  | || ||||  ||||     |||||
Db         119 TTGTCAATAAAACCCTTTCCTTACTCCAGGAGGAAAATGTAGGTA 75

Search completed: September 25, 2003, 14:26:53
Job time : 1653 secs
```

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:59:01 ; Search time 19 Seconds
(without alignments)
794.657 Million cell updates/sec

Title: US-10-049-568-2
Perfect score: 826
Sequence: 1 AQIYSVAIFLGINLAAFIII.........BMLHHSSGGKCGHCRRCHLS 157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	207.5	25.1	1115	2	S40241	G protein-coupled
2	143.5	17.4	695	2	JC1493	follitropin recept
3	143	17.3	925	2	JC2033	G protein-coupled
4	141.5	17.1	695	1	JN0898	follitropin recept
5	140.5	17.0	696	2	A41344	lutropin-choriogon
6	139.5	16.9	692	2	A34548	follitropin recept
7	139.5	16.9	695	2	I45896	follicle stimulati
8	136.5	16.5	694	2	JC4301	follitropin recept
9	134	16.2	700	2	I77463	luteinizing hormon
10	134	16.2	700	2	A49744	lutropin-choriogon
11	134	16.2	700	2	A42395	lutropin receptor
12	134	16.2	814	2	JC7389	thyroid stimulatin
13	132	16.0	793	2	JC7390	thyroid stimulatin
14	131.5	15.9	694	2	JC2237	follitropin recept
15	131	15.9	699	1	QRHUUT	lutropin-choriogon
16	128	15.5	320	2	S17177	probable G protein
17	128	15.5	337	2	S68678	adenosine receptor
18	127	15.4	764	2	JC5643	thyroid stimulatin
19	125.5	15.2	695	1	QRHUFT	follitropin recept
20	124	15.0	696	2	JC7361	follitropin recept
21	124	15.0	764	2	A35956	thyrotropin recept
22	122	14.8	764	2	A40077	thyrotropin recept
23	121	14.6	764	2	I48882	thyrotropin recept
24	119.5	14.5	332	2	A42171	A2-adenosine recep
25	119.5	14.5	378	2	JC5245	G protein-coupled
26	119.5	14.5	417	2	E30341	alpha-1-adrenergic
27	119.5	14.5	517	2	A45121	alpha-1B adrenergi
28	118	14.3	332	2	I48933	adenosine receptor
29	117.5	14.2	515	2	A40491	alpha-1-adrenergic
30	117.5	14.2	764	1	QRHURH	thyrotropin recept
31	117	14.2	326	2	I48096	A1 adenosine recep
32	117	14.2	326	2	A38144	adenosine receptor
33	117	14.2	332	2	JC1229	adenosine receptor
34	115	13.9	326	2	A53005	adenosine receptor
35	114	13.8	466	2	T26458	hypothetical prote
36	113.5	13.7	320	2	A46152	A3 adenosine recep
37	113.5	13.7	392	2	S65693	opioid receptor mu
38	113.5	13.7	400	2	I56553	mu opiate receptor
39	113	13.7	326	2	C30341	G protein-coupled
40	113	13.7	444	2	T27866	hypothetical prote
41	112	13.6	328	2	JN0675	adenosine receptor
42	111	13.4	185	2	I48931	adenosine receptor
43	111	13.4	326	2	A40376	adenosine receptor
44	111	13.4	386	1	S34043	oxytocin receptor
45	110.5	13.4	324	2	S55302	A(1) adenosine rec

ALIGNMENTS

RESULT 1
S40241
G protein-coupled receptor - great pond snail
C;Species: Lymnaea stagnalis (great pond snail)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Jun-2001
C;Accession: S40241
R;Tensen, C.P.; Kesteren, E.R.; Planta, R.J.; Cox, K.; Burke, J.F.; Heerikhuizen, H.
submitted to the EMBL Data Library, June 1993
A;Description: A G protein-coupled receptor with LDL-binding motifs suggests a role
A;Reference number: S40241
A;Accession: S40241
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1115 <TEN>
A;Cross-references: EMBL:Z23104; NID:g438128; PID:g438129
C;Superfamily: great pond snail LDL receptor-related G protein-coupled receptor; LDL
C;Keywords: G protein-coupled receptor; transmembrane protein
F;38-77/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;79-113/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;118-153/Domain: LDL receptor ligand-binding repeat homology <LDLA>
F;158-194/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;195-230/Domain: LDL receptor ligand-binding repeat homology <LDLB>
F;233-267/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;274-316/Domain: LDL receptor ligand-binding repeat homology <LDLC>
F;322-361/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F;367-401/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F;406-440/Domain: LDL receptor ligand-binding repeat homology <LDLD>
F;446-483/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F;488-523/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F;584-607/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;608-631/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;632-655/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;656-679/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;704-727/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;774-797/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

```
  Query Match             25.1%;  Score 207.5;  DB 2;  Length 1115;
  Best Local Similarity   33.1%;  Pred. No. 6.9e-13;
  Matches   51;  Conservative   25;  Mismatches   45;  Indels   33;  Gaps    5;

Qy          4 YSVAIFLGINLAAFIIIVFSYGSMFYSV---HQSAITATEIRNQVKKEMILAKRFFFIVF 60
              ||| :|: :|| :|::|  ||  || ||     :||:    | :|      :   :|:|    ||
Db        938 YSVGVFILLNLLSFVLIASSYLWMF-SVAKKTRSAVRTAESKN----DNAMARRMTLIVM 992

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPF------- 113
              ||  ||:|| |   :||         : :|: :      :|||  ||::|||:| ||
Db        993 TDFCCWVPIIVLGFVSLAGARADDQVYAWIAVFVLPLNSATNPVIYTLSTAPFLGNVRKR 1052

Qy        114 -----KEMIHRF-------------WHNYRQRKS 129
                   |  || |                |:| ::||
Db       1053 ANRFRKSFIHSFTGDTKHSYVDDGTTHSYCEKKS 1086
```

```
RESULT 2
JC1493
follitropin receptor - sheep
N;Alternate names: follicle stimulating hormone receptor
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jan-2000
C;Accession: JC1493; I47080
R;Khan, H.; Yarney, T.A.; Sairam, M.R.
Biochem. Biophys. Res. Commun. 190, 888-894, 1993
A;Title: Cloning of alternately spliced mRNA transcripts coding for variants of ovine te
A;Reference number: JC1493; MUID:93176195; PMID:8439338
A;Accession: JC1493
A;Molecule type: mRNA
A;Residues: 1-695 <KHA>
A;Experimental source: testis
R;Yarney, T.A.; Sairam, M.R.; Khan, H.; Ravindranath, N.; Payne, S.; Seidah, N.G.
Mol. Cell. Endocrinol. 93, 219-226, 1993
A;Title: Molecular cloning and expression of the ovine testicular follicle stimulating h
A;Reference number: I47080; MUID:93351750; PMID:8394255
A;Accession: I47080
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-695 <YAR>
A;Cross-references: GB:L07302; NID:g165884; PIDN:AAA31525.1; PID:g165885
C;Genetics:
A;Gene: FSH-R
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F;191,199/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             17.4%;  Score 143.5;  DB 2;  Length 695;
  Best Local Similarity   27.2%;  Pred. No. 1.4e-06;
  Matches   44;  Conservative   32;  Mismatches   59;  Indels   27;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::: | :|: ||::|   |  :: :|     ||::          :  :|||   ::|
Db        527 SQLYVMSL-LVLNVLAFVVICGCYTHIYLTVRNPNITSS------SSDTKIAKRMAMLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |   ||      |  : :  ::: :  |||  || || : || |:
Db        580 TDFLCMAPISFFAISASLKVPLITVSKSKILLVLFYPINSCANPFLYAIFTRNFRRDFFI 639

Qy        116 MIHRFW------HNYRQRKSMDSKGIRKHMLHHSSGGKCGHC 151
              :: :|             ||  |  :              |:   : |||
Db        640 LLSKFGCYEVQAQTYRSETSFTA---------HNFHPRNGHC 672

RESULT 3
JC2033
G protein-coupled receptor - sea anemone (Anthopleura elegantissima)
C;Species: Anthopleura elegantissima
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 21-Jul-2000
C;Accession: JC2033; S41908
R;Nothacker, H.P.; Grimmelikhuijzen, C.J.P.
Biochem. Biophys. Res. Commun. 197, 1062-1069, 1993
A;Title: Molecular cloning of a novel, putative G protein-coupled receptor from sea anem
A;Reference number: JC2033; MUID:94107299; PMID:8280121
A;Accession: JC2033
A;Molecule type: mRNA
A;Residues: 1-925 <NOT>
A;Cross-references: EMBL:Z28332; NID:g452810; PIDN:CAA82186.1; PID:g452811
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
C;Keywords: G protein-coupled receptor; glycoprotein; neuropeptide; phosphoprotein; tran
F;61-84/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;85-109/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;110-133/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;134-157/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;333-349,350-384,385-419,420-453,454-462/Region: glycine-rich
F;530-551/Domain: transmembrane #status predicted <TM1>
F;562-584/Domain: transmembrane #status predicted <TM2>
F;607-628/Domain: transmembrane #status predicted <TM3>
F;652-673/Domain: transmembrane #status predicted <TM4>
F;692-712/Domain: transmembrane #status predicted <TM5>
F;740-763/Domain: transmembrane #status predicted <TM6>
F;775-795/Domain: transmembrane #status predicted <TM7>
F;61,152,212,435,491,495/Binding site: carbohydrate (Asn) (covalent) #status predicte
F;553,645,855,859,868,894,900/Binding site: phosphate (Ser) (covalent) #status predic
F;649,853/Binding site: phosphate (Thr) (covalent) #status predicted

  Query Match             17.3%;  Score 143;  DB 2;  Length 925;
  Best Local Similarity   33.3%;  Pred. No. 2.1e-06;
  Matches   37;  Conservative   21;  Mismatches   47;  Indels    6;  Gaps    3;

Qy          6 VAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKK-EMILAKRFFFIVFTDAL 64
              ||  | :| |:|| ::: |  | | |    ::  ::     |: :  :|||   :|||| |
Db        694 VAFLLLVNGASFISVMYLYSRMLYVV----VSGGDMEGAPKRNDSKVAKRMAILVFTDML 749

Qy         65 CWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
              || || |        :  |  :  | :  ::: :  |||  || ||   |: ||
Db        750 CWAPIAFFGLLAAFGQTLLTVTQSKILLVFFFPINSICNPFLYAFFTKAFK 800

RESULT 4
JN0898
follitropin receptor precursor - crab-eating macaque
N;Alternate names: follicle-stimulating hormone receptor (FSHR)
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JN0898; S36452
R;Gromoll, J.; Dankbar, B.; Sharma, R.S.; Nieschlag, E.
Biochem. Biophys. Res. Commun. 196, 1066-1072, 1993
A;Title: Molecular cloning of the testicular follicle stimulating hormone receptor of
A;Reference number: JN0898; MUID:94071854; PMID:7504463
A;Accession: JN0898
A;Molecule type: mRNA
A;Residues: 1-695 <GRO>
A;Cross-references: EMBL:X74454; NID:g396801; PIDN:CAA52463.1; PID:g396802
A;Note: the authors translated the codon AGT for residue 488 as Arg
C;Function:
A;Description: receptor that mediates the biochemical effects of follitropin
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotei
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-695/Product: follitropin receptor #status predicted <PFH>
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;367-387/Domain: transmembrane #status predicted <TM1>
F;399-421/Domain: transmembrane #status predicted <TM2>
F;444-465/Domain: transmembrane #status predicted <TM3>
F;486-508/Domain: transmembrane #status predicted <TM4>
F;529-550/Domain: transmembrane #status predicted <TM5>
F;574-597/Domain: transmembrane #status predicted <TM6>
F;609-630/Domain: transmembrane #status predicted <TM7>
F;191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicte
F;596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicte

  Query Match             17.1%;  Score 141.5;  DB 1;  Length 695;
  Best Local Similarity   26.5%;  Pred. No. 2.2e-06;
  Matches   44;  Conservative   32;  Mismatches   63;  Indels   27;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::: | :|: ||::|   |  :: :|     | ::          :  :|||   ::|
Db        527 SQLYVMSL-LVLNVLAFVVICGCYTHIYLTVRNPNIVSS------SSDTRIAKRMAMLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |   ||      |  :    ::: :  |||  || || : |: |:
Db        580 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 639

Qy        116 MIHRFW------HNYRQRKSMDSKGIRKHMLHHSSGGKCGHCRRCH 155
```

```
;    NUMBER OF SEQUENCES:  62
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN DARBY & CUSHMAN
;      STREET:  1615 L STREET, N.W.
;      CITY:  WASHINGTON, D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/741,453A
;      FILING DATE:  19911015
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KOKULIS, PAUL N.
;      REGISTRATION NUMBER:  16773
;      REFERENCE/DOCKET NUMBER:  91913/1107/US/ST
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 861-3000
;      TELEFAX:  (202) 822-0944
;      TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  795 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-741-453A-55

  Query Match             16.6%;  Score 137.5;  DB 3;  Length 795;
  Best Local Similarity   21.9%;  Pred. No. 1.4e-06;
  Matches   43;  Conservative   38;  Mismatches   64;  Indels   51;  Gaps    7;

Qy          1 AQIYSVAIFL-----GINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRF 55
              :|:| ::| :     |:|: ||::|.  |  ::::|    :||        |:  :||:
Db        599 SQVYILSILIRATHCGLNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKM 652

Qy         56 FFIVFTDALCWIPIFVAKPLSLLQVE-------IPGTITSWVVIGYSAINSALNPILYTL 108
                ::|||  |        |:|   :        |  | :  ::: :  :||  || || :
Db        653 AILIFTDFTCMARATHCGPISFFAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAI 712

Qy        109 TTRPFKE---------MIHRFW------HNYRQRK--------------SMDSKGIRK 137
               |: |:           :: ||         ||:::               |  |:   |
Db        713 FTKAFQRDRATHCGFLLLLSRFGCCKRRAELYRRKEFSAYTSNCKNGFPGASKPSQATLK 772

Qy        138 HMLHHSSGGKCGHCRR 153
                |       ||||::
Db        773 --LSTVRATHCGHCQQ 786


RESULT 5
US-08-118-270-18
; Sequence 18, Application US/08118270
; Patent No. 5508384
;  GENERAL INFORMATION:
;    APPLICANT:  Murphy, Randall B.
;    APPLICANT:  Schuster, David I.
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/118,270
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DAIE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  370 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-118-270-18

  Query Match             16.5%;  Score 136;  DB 1;  Length 370;
  Best Local Similarity   28.5%;  Pred. No. 8.1e-07;
  Matches   39;  Conservative   24;  Mismatches   50;  Indels   24;  Gaps    5;

Qy          2 QIYSVAIFLGINLAAFIIIVFSYGSMFYSV--------HQSAITATEIRNQVKKEMILAK 53
              ::| ||     || | ::   | | | ::        |:  :::|: :    :  |  |
Db        178 RVYIVAKRTTKNLEAGVMKEMS-NSKFLTLRIHWSKNFHEDTLSSTKAKGHNPRSSIAVK 236

Qy         54 RFFF-----------IVFTDALCWIPIFVAKPLSLL--QVEIPGTITSWVVIGYSAINSA 100
               | |           ||    |||:| |:| ||  |   :: |  :  |  :||   ||
Db        237 LFKFSREKKAAKTLGIVVGWILCWLPFFIALPLGSLFSTLKPPDAVFKWFWLGY--FNSC 294

Qy        101 LNPILYTLTTRPFKEMI 117
              ||||:|  ::: ||  :
Db        295 LNPIIYPCSSKEFKRAL 311


RESULT 6
PCI-US93-08528-18
; Sequence 18, Application PC/TUS9308528
;  GENERAL INFORMATION:
;    APPLICANT:  New York University
;    TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;    TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;    NUMBER OF SEQUENCES:  348
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/08528
;      FILING DATE:  09-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/943,236
;      FILING DATE:  10-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Kevin G.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  MURPHY=2 PCT
```

```
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-628-5197
;       TELEFAX:  202-737-3528
;       TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  370 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US93-08528-18

  Query Match             16.5%;  Score 136;  DB 5;  Length 370;
  Best Local Similarity   28.5%;  Pred. No. 8.1e-07;
  Matches   39;  Conservative   24;  Mismatches   50;  Indels   24;  Gaps    5;

Qy          2 QIYSVAIFLGINLAAFIIIVFSYGSMFYSV--------HQSAITATEIRNQVKKEMILAK 53
              ::| ||      || | ::    |   | | ::          |:  :::|: :     :   |   |
Db        178 RVYIVAKRTTKNLEAGVMKEMS-NSKFLTLRIHWSKNFHEDTLSSTKAKGHNPRSSIAVK 236

Qy         54 RFFF-----------IVFTDALCWIPIFVAKPLSLL--QVEIPGTITSWVVIGYSAINSA 100
               | |             ||     |||:| |:| ||   |     :: |   :   |   :||    ||
Db        237 LFKFSREKKAAKTLGIVVGWILCWLPFFIALPLGSLFSTLKPPDAVFKWFWLGY--FNSC 294

Qy        101 LNPILYTLTTRPFKEMI 117
              ||||:|   ::: ||   :
Db        295 LNPIIYPCSSKEFKRAL 311


RESULT 7
US-07-757-342D-3
; Sequence 3, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;  INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 700 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-07-757-342D-3

  Query Match             16.2%;  Score 134;  DB 3;  Length 700;
  Best Local Similarity   27.2%;  Pred. No. 2.9e-06;
  Matches   34;  Conservative   31;  Mismatches   48;  Indels   12;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
               :|:| ::| : :|: ||::|   |  ::::|     :||          |:  :||:   ::|
Db        528 SQVYILSILI-LNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIF 580

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              ||  |  || | |    :         |  | :  ::: :   :||  || || : |: |:
Db        581 TDFTCMAPISFFAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQRDFLL 640

Qy        116 MIHRF 120
              :: ||
Db        641 LLSRF 645


RESULT 8
US-07-757-342D-8
; Sequence 8, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;  INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 611 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-07-757-342D-8

  Query Match             15.9%;  Score 131;  DB 3;  Length 611;
  Best Local Similarity   29.6%;  Pred. No. 5.1e-06;
  Matches   34;  Conservative   25;  Mismatches   48;  Indels    8;  Gaps    3;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
               :|:| : | : :|: || ||   |  ::::|     : ||         |:  :||:   ::|
Db        436 SQVYILTILI-LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIF 488

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
              ||  |  || | |    :         |  | :  ::: :   |||  || || : |: |:
Db        489 TDFTCMAPISFFAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 543
```

F;78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;368-389/Domain: transmembrane #status predicted <TM1>
F;400-422/Domain: transmembrane #status predicted <TM2>
F;444-466/Domain: transmembrane #status predicted <TM3>
F;489-511/Domain: transmembrane #status predicted <TM4>
F;530-551/Domain: transmembrane #status predicted <TM5>
F;575-598/Domain: transmembrane #status predicted <TM6>
F;610-631/Domain: transmembrane #status predicted <TM7>
F;103,178,199,295,303,317/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;647,648/Binding site: palmitate (Cys) (covalent) #status predicted
F;681/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

```
  Query Match             16.2%;  Score 134;  DB 2;  Length 700;
  Best Local Similarity   27.2%;  Pred. No. 1.3e-05;
  Matches   34;  Conservative   31;  Mismatches   48;  Indels   12;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::| : :|: ||::|   |  ::::|     :||         |:  :||:   ::|
Db        528 SQVYILSILI-LNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIF 580

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              ||  |  || |  |    :       |   |  :  ::: :   :||   || ||  : |: |:
Db        581 TDFTCMAPISFFAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQRDFLL 640

Qy        116 MIHRF 120
               :: ||
Db        641 LLSRF 645
```

RESULT 11
A42395
lutropin receptor - mouse
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C;Accession: A42395
R;Gudermann, T.; Birnbaumer, M.; Birnbaumer, L.
J. Biol. Chem. 267, 4479-4488, 1992
A;Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenyl
mone receptor expressed in L cells.
A;Reference number: A42395; MUID:92165799; PMID:1311310
A;Accession: A42395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-700 <GUD>
A;Cross-references: GB:M81310; GB:M81318; NID:g198811; PIDN:AAA39432.1; PID:g198812
A;Note: sequence extracted from NCBI backbone (NCBIN:84064, NCBIP:84066)
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: G protein-coupled receptor; transmembrane protein
F;54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

```
  Query Match             16.2%;  Score 134;  DB 2;  Length 700;
  Best Local Similarity   28.0%;  Pred. No. 1.3e-05;
  Matches   35;  Conservative   29;  Mismatches   49;  Indels   12;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::| | :|  ||::|   |  ::::|     :||         |:  :||:   ::|
Db        528 SQVYILSILL-LNAVAFVVICACYVRIYFAVQNPELTAP------NKDTKIAKKMAILIF 580

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLITRPFKE---- 115
              ||  |  || |  |    :       |   |  :  ::: :   :||   || ||  : |: |:
Db        581 TDFTCMAPISFFAISAAFKVPLITVTNSKVLLVLFYPVNSCANPFLYAVFTKAFQRDFFL 640
```

```
Qy        116 MIHRF 120
               :: ||
Db        641 LLSRF 645
```

RESULT 12
JC7389
thyroid stimulating hormone receptor a - salmon
N;Alternate names: thyrotropin receptor a
C;Species: Oncorhynchus sp. (salmon)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 07-Mar-2003
C;Accession: JC7389
R;Oba, Y.; Hirai, T.; Yoshiura, Y.; Kobayashi, T.; Nagahama, Y.
Biochem. Biophys. Res. Commun. 276, 258-263, 2000
A;Title: Cloning, functional characterization, and expression of thyrotropin recepto
A;Reference number: JC7389
A;Contents: Thyroid
A;Accession: JC7389
A;Molecule type: mRNA
A;Residues: 1-814 <OBA>
A;Cross-references: DDBJ:AB030954
C;Comment: This receptor, a transmembrane protein, which mediates the actions of thy
C;Genetics:
A;Gene: tsh-ra
A;Introns: 610/3
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repe
C;Keywords: glycoprotein; hormone receptor; thyroid gland; transmembrane protein

```
  Query Match             16.2%;  Score 134;  DB 2;  Length 814;
  Best Local Similarity   26.8%;  Pred. No. 1.5e-05;
  Matches   44;  Conservative   32;  Mismatches   60;  Indels   28;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              ||:| |:: : :|: ||::|   |  :: :||      :            |:  :|||   ::|
Db        624 AQVYIVSVLI-LNILAFLVICACYIKIYCTVHNPHYRSG------SKDTNIAKRMAVLIF 676

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK----- 114
              || ||  || | |    | :   |  : :  ::: :  :||  || || : |: |:
Db        677 TDFLCMAPISFYAMSAVLDRPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFRGDVFI 736

Qy        115 -----EMIHRFWHNYRQRKSMDSKGIRKHMLHHSSGGKCGHCRR 153
                    :  |    :|  ::: |||          | | |   ||
Db        737 LLSKVGLCQRRAQLFR-GQTVSSKG---------SSGVCHQGRR 770
```

RESULT 13
JC7390
thyroid stimulating hormone receptor b - salmon
N;Alternate names: thyrotropin receptor b
C;Species: Oncorhynchus sp. (salmon)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 07-Mar-2003
C;Accession: JC7390
R;Oba, Y.; Hirai, T.; Yoshiura, Y.; Kobayashi, T.; Nagahama, Y.
Biochem. Biophys. Res. Commun. 276, 258-263, 2000
A;Title: Cloning, functional characterization, and expression of thyrotropin recepto
A;Reference number: JC7389
A;Contents: Thyroid
A;Accession: JC7390
A;Molecule type: mRNA
A;Residues: 1-793 <OBA>
A;Cross-references: DDBJ:AB030955
C;Comment: This receptor, a transmembrane protein, which medaites the actions of thy
C;Genetics:
A;Gene: tsh-rb
A;Introns: 584/3
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repe
C;Keywords: glycoprotein; hormone receptor; thyroid gland; transmembrane protein

```
  Query Match             16.0%;  Score 132;  DB 2;  Length 793;
  Best Local Similarity   27.3%;  Pred. No. 2.4e-05;
  Matches   33;  Conservative   28;  Mismatches   40;  Indels   20;  Gaps    4;
```

```
Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                ||:| ::: : :|: ||::|   |  :: :||       :         |:  :|||    ::|
Db        598 AQVYIISVLI-LNILAFLVICTCYVKIYCAVHNPYYCSG------SKDTNIAKRMAILIF 650

Qy         61 TDALCWIPI-------FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPF 113
                || ||  ||           | :||       |  : :  ::: :  :||  || || : |: |
Db        651 TDFLCMAPISFYAMSAVVDRPL------ITVSNSKILLVLFYPLNSCANPFLYAIFTKAF 704

Qy        114 K 114
                :
Db        705 R 705
```

```
RESULT 14
JC2237
follitropin receptor, testis - horse
N;Alternate names: eFSHR
C;Species: Equus caballus (domestic horse)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 13-Aug-1999
C;Accession: JC2237; JC2370
R;Robert, P.; Amsellem, S.; Christophe, S.; Benifla, J.L.; Bellet, D.; Koman, A.; Bidart
Biochem. Biophys. Res. Commun. 201, 201-207, 1994
A;Title: Cloning and sequencing of the equine testicular follitropin receptor.
A;Reference number: JC2237; MUID:94256980; PMID:8198575
A;Accession: JC2237
A;Molecule type: mRNA
A;Residues: 1-694 <ROB>
A;Cross-references: GB:S70150; NID:g546896; PIDN:AAB30854.1; PID:g546897
A;Experimental source: testis
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: glycoprotein; hormone receptor; transmembrane protein
F;56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;366-386/Domain: transmembrane #status predicted <TM1>
F;398-420/Domain: transmembrane #status predicted <TM2>
F;443-464/Domain: transmembrane #status predicted <TM3>
F;485-507/Domain: transmembrane #status predicted <TM4>
F;528-549/Domain: transmembrane #status predicted <TM5>
F;573-596/Domain: transmembrane #status predicted <TM6>
F;608-629/Domain: transmembrane #status predicted <TM7>
F;191,199,268,293/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             15.9%;  Score 131.5;  DB 2;  Length 694;
  Best Local Similarity   26.5%;  Pred. No. 2.3e-05;
  Matches   43;  Conservative   32;  Mismatches   60;  Indels   27;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::: | :|: ||::|   |   :: :|     | ::          :  :|||    ::|
Db        526 SQLYVMSL-LVLNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTKIAKRMAILIF 578

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
                || ||  || | |    ||      |  : :   ::: :   |||   || ||  : |: |:
Db        579 TDFLCMAPISFFAISASLKVPLITVSKSKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 638

Qy        116 MIHRFW------HNYRQRKSMDSKGIRKHMLHHSSGGKCGHC 151
                :: :|         ||    |   :              | |   : |||
Db        639 LLSKFGCYEMQAQLYRTETSSTA---------HISHPRNGHC 671

RESULT 15
QRHUUT
lutropin-choriogonadotropin receptor precursor - human
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
N;Contains: lutropin-choriogonadotropin receptor precursor long splice form; lutropin-ch
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1993 #sequence_revision 16-Feb-1996 #text_change 22-Jun-1999
C;Accession: A36243; B36243; A23728; B36120
R;Minegish, T.; Nakamura, K.; Takakura, Y.; Miyamoto, K.; Hasegawa, Y.; Ibuki, Y.; Igara
```

```
Biochem. Biophys. Res. Commun. 172, 1049-1054, 1990
A;Title: Cloning and sequencing of human LH/hCG receptor cDNA.
A;Reference number: A36243; MUID:91058534; PMID:2244890
A;Accession: A36243
A;Molecule type: mRNA
A;Residues: 1-699 <MIN>
A;Cross-references: GB:M63108; NID:g187135; PIDN:AAA59515.1; PID:g187136
A;Experimental source: ovary
A;Accession: B36243
A;Molecule type: mRNA
A;Residues: 1-226,290-699 <MI2>
R;Jia, X.C.; Oikawa, M.; Bo, M.; Tanaka, T.; Ny, T.; Boime, I.; Hsueh, A.J.W.
Mol. Endocrinol. 5, 759-768, 1991
A;Title: Expression of human luteinizing hormone (LH) receptor: interaction with LH a
A;Reference number: A23728; MUID:92017881; PMID:1922095
A;Accession: A23728
A;Molecule type: mRNA
A;Residues: 1-123,'R',125-311,'N',313-699 <JIA>
A;Cross-references: GB:S57793
A;Experimental source: ovary granulosa and luteal cells; thyroid
A;Note: 312-Ser was also found; the authors translated the codon UAC for residue 546
A;Note: mRNA from the thyroid is incompletely processed compared with that from the o
R;Frazier, A.L.; Robbins, L.S.; Stork, P.J.; Sprengel, R.; Segaloff, D.L.; Cone, R.D.
Mol. Endocrinol. 4, 1264-1276, 1990
A;Title: Isolation of TSH and LH/CG receptor cDNAs from human thyroid: regulation by
A;Reference number: A36120; MUID:91155962; PMID:2293030
A;Accession: B36120
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-6,'P',8-18,'A',20-26,'R',29-43,'A',45,'A',47,'RPS',52-67,'S',69-123,'R'
A;Cross-references: GB:M73746; NID:g903745; PIDN:AAA70231.1; PID:g903746
A;Experimental source: thyroid
R;Zhu, H.; Wang, H.; Ascoli, M.
Mol. Endocrinol. 9, 141-150, 1995
A;Title: The lutropin/choriogonadotropin receptor is palmitoylated at intracellular c
A;Reference number: A57564; MUID:95295727; PMID:7776964
A;Contents: annotation; mutation analysis of palmitate binding sites
C;Genetics:
A;Gene: GDB:LHCGR
A;Cross-references: GDB:125260; OMIM:152790
A;Map position: 2p21-2p21
A;Introns: 226/3
A;Note: the exact position of the intron cannot be determined from the experimental d
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone r
F;1-699/Product: lutropin-choriogonadotropin receptor precursor, long splice form #st
F;1-226,290-699/Product: lutropin-choriogonadotropin receptor precursor, short splice
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-644/Product: lutropin-choriogonadotropin receptor #status predicted <MAT>
F;23-363/Domain: extracellular hormone binding #status predicted <EHB>
F;50-73/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;74-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;124-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;149-173/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;176-197/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;198-222/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;364-385/Domain: transmembrane #status predicted <TM1>
F;396-418/Domain: transmembrane #status predicted <TM2>
F;441-462/Domain: transmembrane #status predicted <TM3>
F;483-505/Domain: transmembrane #status predicted <TM4>
F;526-547/Domain: transmembrane #status predicted <TM5>
F;571-594/Domain: transmembrane #status predicted <TM6>
F;606-627/Domain: transmembrane #status predicted <TM7>
F;645-699/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;99,174,195,291,299,313/Binding site: carbohydrate (Asn) (covalent) #status predicte
F;561,677/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status pred
F;643,644/Binding site: palmitate (Cys) (covalent) #status experimental
F;670/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicte

  Query Match             15.9%;  Score 131;  DB 1;  Length 699;
  Best Local Similarity   29.6%;  Pred. No. 2.6e-05;
  Matches   34;  Conservative   25;  Mismatches   48;  Indels    8;  Gaps    3;
```

```
RESULT 14
US-08-795-876-2
; Sequence 2, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  436 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-795-876-2

  Query Match          15.2%;  Score 125.5;  DB 4;  Length 436;
  Best Local Similarity   26.1%;  Pred. No. 1.4e-05;
  Matches   43;  Conservative   30;  Mismatches   59;  Indels   33;  Gaps    6;

Qy          1 AQTYSVAIFLGILNLAAFIIIVFSYGSMFYSVHQSAITAYEIRNQVKKEMILAKRFFFIVF 60
              :|:| ::: | :|: ||::|   |  :: :|    | ::        :  :|||   ::|
Db        268 SQLVVMSL-LVLWVLAFVVICGCYIHIYILTVRNPNIVSS------SSDTRIAKRMAMLIF 320

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFK----- 114
               || ||  || | |    ||    |  :    ::: :  |||  || || : |: ||:
Db        321 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFHPINSCANPFLYAITKNFRRDFPI 380

Qy         115 --------EMIHRFWHNYRQRKSMDSKGIRKHMLHSSSGGKCGHC 151
                       ||  :    ||   |              ::  : |||
Db         381 LLSKCGCYEMQAQI---YRTETSSTV---------HWTHPRNGHC 413


RESULT 15
US-08-487-886-2
; Sequence 2, Application US/08487886
; Patent No. 5744448
;  GENERAL INFORMATION:
;    APPLICANT:  Kelton, Christie Ann
;    APPLICANT:  Schweickhardt, Rene Lynn
;    APPLICANT:  Cheng, Shirley Yui Yen
;    APPLICANT:  Nugent, No. 5744448een Patrice
;    TITLE OF INVENTION:  Human Follicle Stimulating
;    TITLE OF INVENTION:  Hormone Receptor
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Stephan P. Williams,
;      ADDRESSEE:  Ares-Serono, Inc.
;      STREET:  Exchange Place, 37th floor
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" diskette, 1.44 MB, high density
;      COMPUTER:  IBM PS/2, model 55 SX
;      OPERATING SYSTEM:  MS-DOS version 4.0
;      SOFTWARE:  VAX/VMS Msgsll via Kermit to IBM MS-DOS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,886
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/670,085
;      FILING DATE:  15-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Stephan P.
;      REGISTRATION NUMBER:  28546
;      REFERENCE/DOCKET NUMBER:  US/252
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 723-1300
;      TELEFAX:  (617) 723-8923
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  695
;      TYPE:  amino acid
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  signal sequence
;      LOCATION:  -17 to -1
;      IDENTIFICATION METHOD:  hydrophobic
;    FEATURE:
;      NAME/KEY:  putative amino-terminal extracellular domain
;      LOCATION:  1 to 349
;      IDENTIFICATION METHOD:  similarity with other
;      IDENTIFICATION METHOD:  dimeric glycoprotein receptor extracellular
;      IDENTIFICATION METHOD:  domains, hydrophilic
;    FEATURE:
;      NAME/KEY:  transmembrane domain
;      LOCATION:  350 to 613
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane domains
;    FEATURE:
;      NAME/KEY:  putative transmembrane region I
;      LOCATION:  350 to 370
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region II
;      LOCATION:  382 to 404
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region III
;      LOCATION:  427 to 448
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region IV
;      LOCATION:  469 to 491
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
```

```
;      NAME/KEY:  putative transmembrane region V
;      LOCATION:  512 to 533
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region VI
;      LOCATION:  557 to 580
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative transmembrane region VII
;      LOCATION:  592 to 613
;      IDENTIFICATION METHOD:  similarity to other G
;      IDENTIFICATION METHOD:  protein-coupled receptor transmembrane regions,
;      IDENTIFICATION METHOD:  hydrophobic, about 20-23 amino acids in length
;    FEATURE:
;      NAME/KEY:  putative carboxy-terminal intracellular
;      NAME/KEY:  domain
;      LOCATION:  614 to 678
US-08-487-886-2

  Query Match            15.2%;  Score 125.5;  DB 1;  Length 695;
  Best Local Similarity   26.1%;  Pred. No. 2.4e-05;
  Matches   43;  Conservative   30;  Mismatches   59;  Indels   33;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| ::: | :|: ||::|   |  :: :|    | ::         :  :|||   ::|
Db        527 SQLYVMSL-LVLNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTRIAKRMAMLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK----- 114
              || ||  || | |   ||     |  :    ::: :  |||  || || : |: |:
Db        580 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFHPINSCANPFLYAIFTKNFRRDFFI 639

Qy        115 --------EMIHRFWHNYRQRKSMDSKGIRKHMLHHSSGGKCGHC 151
                      ||  :    ||   |           |::  : |||
Db        640 LLSKCGCYEMQAQI---YRTETSSTV---------HNTHPRNGHC 672

Search completed: September 25, 2003, 15:01:14
Job time : 31 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 25, 2003, 15:00:41 ; Search time 64 Seconds
(without alignments)
371.175 Million cell updates/sec

Title: US-10-049-568-2
Perfect score: 826
Sequence: 1 AQIYSVAIFLGINLAAFIII.....HMLHRSSGGKCGHCRRCRLS 157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	638	77.2	188	15	US-10-073-885-69	Sequence 69, Appl
2	638	77.2	396	10	US-09-895-686-4	Sequence 4, Appli
3	638	77.2	757	10	US-09-928-175-24	Sequence 24, Appl
4	638	77.2	757	15	US-10-225-567A-623	Sequence 623, App
5	632	76.5	176	9	US-09-761-853-557	Sequence 557, App
6	632	76.5	176	10	US-09-764-877-2009	Sequence 2009, Ap
7	632	76.5	176	11	US-09-989-442-105	Sequence 105, App
8	632	76.5	176	12	US-09-764-886-52	Sequence 52, Appl
9	632	76.5	176	15	US-10-073-865-84	Sequence 84, Appl
10	632	76.5	176	15	US-10-103-313-373	Sequence 373, App
11	632	76.5	176	15	US-10-073-885-96	Sequence 96, Appl
12	409	49.5	718	10	US-09-928-175-21	Sequence 21, Appl
13	409	49.5	737	10	US-09-928-175-20	Sequence 20, Appl
14	409	49.5	737	15	US-10-229-735-3	Sequence 3, Appli
15	392.5	47.5	355	12	US-10-321-807-18	Sequence 18, Appl
16	392.5	47.5	646	10	US-09-928-175-13	Sequence 13, Appl
17	392.5	47.5	682	10	US-09-928-175-12	Sequence 12, Appl
18	392.5	47.5	694	10	US-09-928-175-8	Sequence 8, Appli
19	392.5	47.5	713	11	US-09-965-536A-6	Sequence 6, Appli
20	392.5	47.5	718	10	US-09-928-175-3	Sequence 3, Appli
21	392.5	47.5	730	10	US-09-928-175-7	Sequence 7, Appli
22	392.5	47.5	737	11	US-09-965-536A-2	Sequence 2, Appli
23	392.5	47.5	754	10	US-09-928-175-2	Sequence 2, Appli
24	392.5	47.5	754	15	US-10-229-735-1	Sequence 1, Appli
25	392.5	47.5	754	15	US-10-222-668-2	Sequence 2, Appli
26	349	42.3	107	11	US-09-930-312-2	Sequence 2, Appli
27	244	29.5	327	11	US-09-782-974C-66	Sequence 66, Appl
28	224.5	27.2	334	15	US-10-270-333-21	Sequence 21, Appl
29	224.5	27.2	722	15	US-10-222-668-4	Sequence 4, Appli
30	212.5	25.7	359	15	US-10-270-333-27	Sequence 27, Appl
31	207.5	25.1	1115	11	US-09-965-536A-10	Sequence 10, Appl
32	207.5	25.1	1115	15	US-10-222-668-3	Sequence 3, Appli
33	139.5	16.9	675	9	US-09-877-804-7	Sequence 7, Appli
34	139.5	16.9	688	11	US-09-965-536A-12	Sequence 12, Appl
35	139.5	16.9	692	9	US-09-877-804-6	Sequence 6, Appli
36	139.5	16.9	692	11	US-09-965-536A-11	Sequence 11, Appl
37	134	16.2	674	9	US-09-877-804-3	Sequence 3, Appli
38	134	16.2	700	9	US-09-877-804-2	Sequence 2, Appli
39	131	15.9	458	9	US-09-862-767A-5	Sequence 5, Appli
40	131	15.9	693	11	US-09-965-536A-14	Sequence 14, Appl
41	131	15.9	699	10	US-09-804-626-6	Sequence 6, Appli
42	131	15.9	699	15	US-10-225-567A-150	Sequence 150, App
43	130	15.7	861	9	US-09-804-551B-20	Sequence 20, Appl
44	129.5	15.7	695	12	US-10-349-836A-30	Sequence 30, Appl
45	129	15.6	687	11	US-09-965-536A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-073-885-69
; Sequence 69, Application US/10073885
; Publication No. US20030096346A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ03C1
; CURRENT APPLICATION NUMBER: US/10/073,885
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-073-885-69

Query Match 77.2%; Score 638; DB 15; Length 188;
Best Local Similarity 94.8%; Pred. No. 2.1e-62;
Matches 127; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 AQIYSVAIFLGINLAAFIIIVPSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFPFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 AQIYSVAIFLGINLAAFIIIVPSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFPFIVF 66

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRP 120
              |||||||||| | ||||||||||||||||     ||||||||||||||||||||||||||
Db         67 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF 126

Qy        121 WENYRQRKSMDSKG 134
              |:||||||||||||
Db        127 WYNYRQRKSMDSKG 140
```

RESULT 2
US-09-895-686-4

```
; Sequence 4, Application US/09895686
; Patent No. US20020106655A1
; GENERAL INFORMATION:
;  APPLICANT: Bandman, Olga
;  APPLICANT:  Lal, Preeti
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Baughn, Mariah R.
;  TITLE OF INVENTION: HUMAN GPCR PROTEINS
;  FILE REFERENCE: PC-0044 CIP
;  CURRENT APPLICATION NUMBER: US/09/895,686
;  CURRENT FILING DATE:  2001-06-28
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PERL Program
; SEQ ID NO 4
;    LENGTH: 396
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. US20020106655A1 2488822CD1
US-09-895-686-4

  Query Match             77.2%;  Score 638;  DB 10;  Length 396;
  Best Local Similarity   94.8%;  Pred. No. 5.4e-62;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 274

Qy         61 TDALCWIPIPVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||    ||||||||||||||||||||||||||
Db        275 TDALCWIPIPVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF 334

Qy        121 WHNYRQRKSMDSKG 134
              |:|||||||||||
Db        335 WYNYRQRKSMDSKG 348


RESULT 3
US-09-928-175-24
; Sequence 24, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Passty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, Mc. US20020123618A1ma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;    LENGTH: 757
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-09-928-175-24

  Query Match             77.2%;  Score 638;  DB 10;  Length 757;
  Best Local Similarity   94.8%;  Pred. No. 1.2e-61;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 635

Qy         61 TDALCWIPIPVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||    ||||||||||||||||||||||||||
Db        636 TDALCWIPIPVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF 695

Qy        121 WHNYRQRKSMDSKG 134
              |:|||||||||||
Db        696 WYNYRQRKSMDSKG 709


RESULT 4
US-10-225-567A-623
; Sequence 623, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burner, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPT
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 623
;    LENGTH: 757
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-225-567A-623

  Query Match             77.2%;  Score 638;  DB 15;  Length 757;
  Best Local Similarity   94.8%;  Pred. No. 1.2e-61;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 635

Qy         61 TDALCWIPIPVVKFLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||    ||||||||||||||||||||||||||
Db        636 TDALCWIPIPVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTTRPFKEMIHRF 695

Qy        121 WHNYRQRKSMDSKG 134
              |:|||||||||||
Db        696 WYNYRQRKSMDSKG 709


RESULT 5
US-09-764-853-557
; Sequence 557, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJ506
;  CURRENT APPLICATION NUMBER: US/09/764,853
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 939
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 557
;    LENGTH: 176
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: SITE
;    LOCATION: (133)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-557

  Query Match             76.5%;  Score 632;  DB 9;  Length 176;
  Best Local Similarity   94.0%;  Pred. No. 9e-62;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;
```

```
                :: :|          ||    |  :          |:|  : |||   |
Db        640 LLSKFGCYEMQAQIYRTETSSTA---------HNSHPRNGHCSSAH 676
```

RESULT 5
A41344
lutropin-choriogonadotropin receptor precursor (splice form A) - pig
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 21-Jan-2000
C;Accession: A41344
R;Loosfelt, H.; Misrahi, M.; Atger, M.; Salesse, R.; Vu Hai-Luu Thi, M.T.; Jolivet, A.;
Science 245, 525-528, 1989
A;Title: Cloning and sequencing of porcine LH-hCG receptor cDNA: variants lacking transm
A;Reference number: A41344; MUID:89332517; PMID:2502844
A;Accession: A41344
A;Molecule type: mRNA
A;Residues: 1-696 <LOO>
A;Cross-references: GB:M29525; NID:g164528; PIDN:AAA31062.1; PID:g164529
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; membrane prc
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-696/Product: lutropin-choriogonadotropic hormone receptor #status predicted <MAT>
F;99-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

```
  Query Match             17.0%;  Score 140.5;  DB 2;  Length 696;
  Best Local Similarity   27.2%;  Pred. No. 2.8e-06;
  Matches   41;  Conservative   26;  Mismatches   55;  Indels   29;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| : | : :|: |||||   |  ::::|      : ||         |:  :||:    ::|
Db        524 SQVYILTILI-LNVVAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIF 576

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHR 119
                || |  || | |  :|    |  | :  ::: :  :||  || || : |: |:
Db        577 TDFTCMAPISFFAISAALKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFR----- 631

Qy        120 FWHNYRQRKSMDSKGIRKHMLHHSSGGKCGH 150
                                |  |  |  | | |
Db        632 ----------------RDFFLLLSKSGCCKH 646
```

RESULT 6
A34548
follitropin receptor precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
C;Accession: A34548; A41729
R;Sprengel, R.; Braun, T.; Nikolics, K.; Segaloff, D.L.; Seeburg, P.H.
Mol. Endocrinol. 4, 525-530, 1990
A;Title: The testicular receptor for follicle stimulating hormone: structure and functic
A;Reference number: A34548; MUID:91125358; PMID:2126341
A;Accession: A34548
A;Molecule type: mRNA
A;Residues: 1-692 <SPR>
A;Cross-references: GB:L02842; NID:g204183; PIDN:AAA41175.1; PID:g204184
R;Heckert, L.L.; Daley, I.J.; Griswold, M.D.
Mol. Endocrinol. 6, 70-80, 1992
A;Title: Structural organization of the follicle-stimulating hormone receptor gene.
A;Reference number: A41729; MUID:92149579; PMID:1738373
A;Accession: A41729
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-692 <HEC>
A;Cross-references: GB:S81198; NID:g245344; PIDN:AAB21415.1; PID:g245345
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:81117, NCBIN:81119, NCBIN:81121, NC
R;Davis, D.; Liu, X.; Segaloff, D.L.
Mol. Endocrinol. 9, 159-170, 1995
A;Title: Identification of the sites of N-linked glycosylation on the follicle-stimulati
A;Reference number: A57562; MUID:95295729; PMID:7776966
A;Contents: annotation; glycosylation sites
C;Function:
A;Description: receptor that mediates the biochemical effects of follitropin
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone r
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-692/Product: follitropin receptor #status predicted <MAT>
F;16-366/Domain: extracellular hormone binding #status predicted <EHB>
F;56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;367-387/Domain: transmembrane #status predicted <TM1>
F;398-421/Domain: transmembrane #status predicted <TM2>
F;443-465/Domain: transmembrane #status predicted <TM3>
F;486-508/Domain: transmembrane #status predicted <TM4>
F;529-550/Domain: transmembrane #status predicted <TM5>
F;574-597/Domain: transmembrane #status predicted <TM6>
F;609-630/Domain: transmembrane #status predicted <TM7>
F;191,199,293/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;554/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicte
F;595/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicte

```
  Query Match             16.9%;  Score 139.5;  DB 2;  Length 692;
  Best Local Similarity   27.3%;  Pred. No. 3.6e-06;
  Matches   47;  Conservative   30;  Mismatches   60;  Indels   35;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| :|: | :|: ||::|    |  :: :|     | ::          :  :|||   ::|
Db        526 SQLYVMAL-LVLNVLAFVVICGCYTHIYLTVRNPTIVSS------SSDTKIAKRMATLIF 578

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
                || ||  || | |   ||        |  :     ::: :  |||  || || : |: |:
Db        579 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 638

Qy        116 MIHRF-----------------WHNYRQRKSMDSKGIRKH------MLHHSS 144
                :: :|                 ||:  |||   |     |           |:|||
Db        639 LLSKFGCYEMQAQIYRTETSSATHNFHARKSHCSSAPRVTNSYVLVPLNHSS 690
```

RESULT 7
I45896
follicle stimulating hormone receptor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jan-2000
C;Accession: I45896
R;Houde, A.; Lambert, A.; Saumande, J.; Silversides, D.W.; Lussier, J.G.
Mol. Reprod. Dev. 39, 127-135, 1994
A;Title: Structure of the bovine follicle-stimulating hormone receptor complementary
A;Reference number: I45896; MUID:95127199; PMID:7826612
A;Accession: I45896
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-695 <HOU>
A;Cross-references: GB:L22319; NID:g404671; PIDN:AAC37324.1; PID:g404672
C;Genetics:
A;Gene: FSHR
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

```
  Query Match             16.9%;  Score 139.5;  DB 2;  Length 695;
  Best Local Similarity   26.5%;  Pred. No. 3.6e-06;
  Matches   43;  Conservative   33;  Mismatches   59;  Indels   27;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::: | :|: ||::|    |  :: :|     ||::          :  :|||   ::|
Db        527 SQLYVMSL-LVLNVLAFVVICGCYTHIYLTVRNPNITSS------SSDTKIAKRMAMLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
                || ||  || | |   ||        |  :  :  ::: :  |||  || || : |: |:
Db        580 TDFLCMAPISFFAISASLKVPLITVSKSKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 639
```

```
Qy        116 MIHRFW------HNYRQRKSMDSKGIRKHMLHHSSGGKCGHC 151
              :: :|       ||  |  :          |:   : |||
Db        640 LLSKFGCYEVQAQTYRSETSSTA---------HNFHPRNGHC 672
```

RESULT 8
JC4301
follitropin receptor - pig
N;Alternate names: follicle-stimulating hormone receptor
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jan-2000
C;Accession: JC4301
R;Remy, J.J.; Lahbib-Mansais, Y.; Yerle, M.; Bozon, V.; Couture, L.; Pajot, E.; Grebert,
Gene 163, 257-261, 1995
A;Title: The porcine follitropin receptor: cDNA cloning, functional expressionand chromo
A;Reference number: JC4301; MUID:96011644; PMID:7590277
A;Accession: JC4301
A;Molecule type: mRNA
A;Residues: 1-694 <REM>
A;Cross-references: GB:L31966
A;Experimental source: ovarian granulosa cells
C;Comment: This receptor belongs to the family of the G-protein coupled receptors. It ha
ermatogenesis in male and oogenesis in female.
C;Genetics:
A;Gene: fshr
A;Map position: 3 q2.2-q2.3
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F;1-365/Domain: follicle-stimulating hormone binding #status predicted <HOB>
F;70-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F;366-388/Domain: transmembrane #status predicted <TM1>
F;398-420/Domain: transmembrane #status predicted <TM2>
F;443-464/Domain: transmembrane #status predicted <TM3>
F;485-507/Domain: transmembrane #status predicted <TM4>
F;528-549/Domain: transmembrane #status predicted <TM5>
F;573-596/Domain: transmembrane #status predicted <TM6>
F;608-629/Domain: transmembrane #status predicted <TM7>

```
  Query Match             16.5%;  Score 136.5;  DB 2;  Length 694;
  Best Local Similarity   26.8%;  Pred. No. 7.2e-06;
  Matches   42;  Conservative   36;  Mismatches   62;  Indels   17;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| |:: | :|: ||::|   |  :: :|    | ::         :  :|||   ::|
Db        526 SQLYVVSL-LVLNVLAFVVICGCYIHIYLTVRNPNIMSS------SSDTKIAKRMAMLIF 578

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || || :|| |  |     ||     |  :     ::: :  |||  || || : |: |:
Db        579 TDFLCMVPISFFAISASLKVPLITVSKLKILLVLFYPINSCANPFLYAIFTKNFRRDVFI 638

Qy        116 MIHRFWHNYRQRKSMDSKGI-RKHMLHHSSGGKCGHC 151
              :: :|     | ::  :: :   | :|  :     |||
Db        639 LLSKFGCYEMQAQTYRTENLSTAHNIHPRN----GHC 671
```

RESULT 9
I77463
luteinizing hormone/chorionic gonadotropin receptor - rat
C;Species: Rattus sp. (rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C;Accession: I77463
R;Aatsinki, J.T.; Pietila, E.M.; Lakkakorpi, J.T.; Rajaniemi, H.J.
Mol. Cell. Endocrinol. 84, 127-135, 1992
A;Title: Expression of the LH/CG receptor gene in rat ovarian tissue is regulated by an
A;Reference number: I57668; MUID:92347604; PMID:1353463
A;Accession: I77463
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-700 <RES>
A;Cross-references: GB:S40803; NID:g252163; PIDN:AAB22680.1; PID:g252164
C;Genetics:
A;Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

```
  Query Match             16.2%;  Score 134;  DB 2;  Length 700;
  Best Local Similarity   27.2%;  Pred. No. 1.3e-05;
  Matches   34;  Conservative   31;  Mismatches   48;  Indels   12;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| ::| : :|: ||::|   |  ::::|     :||          |:  :||:   ::|
Db        528 SQVYILSILI-LNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIF 580

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              ||  |  || | |    :        |  | :  ::: :  :||  || || : |: |:
Db        581 TDFTCMAPISFFAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQRDFLL 640

Qy        116 MIHRF 120
              :: ||
Db        641 LLSRF 645
```

RESULT 10
A49744
lutropin-choriogonadotropin receptor precursor - rat
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-May-1994 #sequence_revision 13-Sep-1998 #text_change 13-Aug-1999
C;Accession: A49744; A40545; A41343; A61453; A32460
R;Koo, Y.B.; Ji, I.; Slaughter, R.G.; Ji, T.H.
Endocrinology 128, 2297-2308, 1991
A;Title: Structure of the luteinizing hormone receptor gene and multiple exons of the
A;Reference number: A49744; MUID:91209270; PMID:2019252
A;Accession: A49744.
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-700 <KOO>
A;Cross-references: GB:M68917
A;Note: authors translated the codon CAA for residue 307 as Glu, AAC for residue 355
R;Tsai-Morris, C.H.; Buczko, E.; Wang, W.; Xie, X.Z.; Dufau, M.L.
J. Biol. Chem. 266, 11355-11359, 1991
A;Title: Structural organization of the rat luteinizing hormone (LH) receptor gene.
A;Reference number: A40545; MUID:91250455; PMID:2040640
A;Accession: A40545
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-320 <TSA>
A;Cross-references: GB:M63918; GB:M63919; GB:M63920; GB:M63921; GB:M63922; GB:M63923;
R;McFarland, K.C.; Sprengel, R.; Phillips, H.S.; Koehler, M.; Rosemblit, N.; Nikolics
Science 245, 494-499, 1989
A;Title: Lutropin-choriogonadotropin receptor: an unusual member of the G protein-cou
A;Reference number: A41343; MUID:89332512; PMID:2502842
A;Accession: A41343
A;Molecule type: mRNA
A;Residues: 1-700 <MCF>
A;Cross-references: GB:M26199; NID:g205178; PIDN:AAA41528.1; PID:g205179
R;Dufau, M.L.; Minegishi, T.; Buczko, E.S.; Delgado, C.J.; Zhang, R.
J. Steroid Biochem. 33, 715-720, 1989
A;Title: Characterization and structure of ovarian and testicular LH/hCG receptors.
A;Reference number: A61453; MUID:90097014; PMID:2601325
A;Accession: A61453
A;Status: preliminary
A;Molecule type: protein
A;Residues: 27-33,'X',35-37,'X',39,'X',41-44 <DUF>
R;Roche, P.C.; Ryan, R.J.
J. Biol. Chem. 264, 4636-4641, 1989
A;Title: Purification, characterization, and amino-terminal sequence of rat ovarian r
A;Reference number: A32460; MUID:89174723; PMID:2925659
A;Accession: A32460
A;Molecule type: protein
A;Residues: 27-32,'LX',35-37 <ROC>
C;Genetics:
A;Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repea
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone r
F;54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

```
;   NAME/KEY: SITE
;   LOCATION: (133)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-886-52

  Query Match              76.5%;  Score 632;  DB 12;  Length 176;
  Best Local Similarity    94.0%;  Pred. No. 9e-62;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPLLYTLLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPLLYTLLTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSXG 134
              |:|||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 9
US-10-073-865-84
; Sequence 84, Application US/10073865
; Publication No. US20030044904A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ09C1
;  CURRENT APPLICATION NUMBER: US/10/073,865
;  CURRENT FILING DATE:  2002-02-14
;  Prior Application removed - See file Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 154
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
;   LENGTH: 176
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (133)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-073-865-84

  Query Match              76.5%;  Score 632;  DB 15;  Length 176;
  Best Local Similarity    94.0%;  Pred. No. 9e-62;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPLLYTLLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPLLYTLLTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSXG 134
              |:|||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 10
US-10-103-313-373
; Sequence 373, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ07C1
;  CURRENT APPLICATION NUMBER: US/10/103,313
;  CURRENT FILING DATE:  2002-03-12
;  NUMBER OF SEQ ID NOS: 653
;       Prior Application removed - See File Wrapper or Palm
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 373
;   LENGTH: 176
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (133)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-103-313-373

  Query Match              76.5%;  Score 632;  DB 15;  Length 176;
  Best Local Similarity    94.0%;  Pred. No. 9e-62;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPLLYTLLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPLLYTLLTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSXG 134
              |:|||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 11
US-10-073-885-96
; Sequence 96, Application US/10073885
; Publication No. US20030096346A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ03C1
;  CURRENT APPLICATION NUMBER: US/10/073,885
;  CURRENT FILING DATE:  2002-02-14
;  Prior Application removed - See file Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 96
;   LENGTH: 176
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (133)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-073-885-96

  Query Match              76.5%;  Score 632;  DB 15;  Length 176;
  Best Local Similarity    94.0%;  Pred. No. 9e-62;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPLLYTLLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPLLYTLLTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSXG 134
              |:|||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 12
US-09-928-175-21
; Sequence 21, Application US/09928175
; Patent No. US20020123618A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618Alma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;    LENGTH: 718
;    TYPE: PRT
;    ORGANISM: Mus musculus
US-09-928-175-21

  Query Match             49.5%;  Score 409;  DB 10;  Length 718;
  Best Local Similarity   58.9%;  Pred. No. 2e-36;
  Matches   73;  Conservative   27;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          4 YSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKEMILAKRFFPIVFTDA 63
              ||: |||||:|| ||:+|| || :+|| |:|::|:    |:|: : ||: :| ||||||:||
Db        553 YSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFFPIVFSDA 612

Qy         64 LCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRFWHN 123
              :|||||:|| | |||||||||||||||:|: :  :|||||||||||  ||: : :  |
Db        613 ICWIPVFVVKILSLLQVEIPGTITSWIVVPFLFVNSALNPILYTLTTSFFKDKLKQLLHK 672

Qy        124 YRQR 127
                :|::
Db        673 HRRK 676


RESULT 13
US-09-928-175-20
; Sequence 20, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
;  APPLICANT: Paszty, Christopher J.
;  APPLICANT:  Gong, Jianhua
;  APPLICANT:  Daugherty, Betsy
;  APPLICANT:  Rogers, No. US20020123618Alma
;  TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
;  TITLE OF INVENTION:  Uses Thereof
;  FILE REFERENCE: 00-1229
;  CURRENT APPLICATION NUMBER: US/09/928,175
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: 60/224,455
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;    LENGTH: 737
;    TYPE: PRT
;    ORGANISM: Mus musculus
US-09-928-175-20

  Query Match             49.5%;  Score 409;  DB 10;  Length 737;
  Best Local Similarity   58.9%;  Pred. No. 2e-36;
  Matches   73;  Conservative   27;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          4 YSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKEMILAKRFFPIVFTDA 63
              ||: |||||:|| ||:+|| || :|| |:|:::|:    |:|: : ||: :| ||||||:||
Db        572 YSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFFPIVFSDA 631

Qy         64 LCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRFWHN 123
              :|||:|:|| | |||||||||||||||:|: :  :|||||||||||  ||: : :  |
Db        632 ICWIPVFVVKILSLLQVEIPGTITSWIVVPFLFVNSALNPILYTLTTSFFKDKLKQLLHK 691

Qy        124 YRQR 127
                :|::
Db        692 HRRK 695


RESULT 14
US-10-229-735-3
; Sequence 3, Application US/10229735
; Publication No. US20030082650A1
; GENERAL INFORMATION:
;  APPLICANT: Baylor College of Medicine
;  APPLICANT:  Agoulnik, Alexander I.
;  TITLE OF INVENTION: The GR8AT Gene and Protein
;  FILE REFERENCE: 7572/73263
;  CURRENT APPLICATION NUMBER: US/10/229,735
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/315,696
;  PRIOR FILING DATE: 2001-08-30
;  PRIOR APPLICATION NUMBER: 60/351,432
;  PRIOR FILING DATE: 2002-01-28
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;    LENGTH: 737
;    TYPE: PRT
;    ORGANISM: Mus musculus
US-10-229-735-3

  Query Match             49.5%;  Score 409;  DB 15;  Length 737;
  Best Local Similarity   58.9%;  Pred. No. 2e-36;
  Matches   73;  Conservative   27;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          4 YSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKEMILAKRFFPIVFTDA 63
              ||: |||||:|| ||:+|| || :|| |:|:::|:    |:|: : ||: :| ||||||:||
Db        572 YSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFFPIVFSDA 631

Qy         64 LCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRFWHN 123
              :|||:|:|| | |||||||||||||||:|: :  :|||||||||||  ||: : :  |
Db        632 ICWIPVFVVKILSLLQVEIPGTITSWIVVPFLFVNSALNPILYTLTTSFFKDKLKQLLHK 691

Qy        124 YRQR 127
                :|::
Db        692 HRRK 695


RESULT 15
US-10-321-807-18
; Sequence 18, Application US/10321807
; Publication No. US20030166148A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Rupong
;  APPLICANT:  Dang, Huong T.
;  APPLICANT:  Lowitz, Kevin P.
;  TITLE OF INVENTION: No. US20030166148A1-Endogenous, Constitutively Activated Human
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: AREN0086
;  CURRENT APPLICATION NUMBER: US/10/321,807
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: US/09/714,008
;  PRIOR FILING DATE: 2000-11-16
;  PRIOR APPLICATION NUMBER: 09/170,496
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: PCT/US99/21938
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: 60/166,088
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/166,099
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/166,369
;  PRIOR FILING DATE: 1999-11-17
;  PRIOR APPLICATION NUMBER: 60/171,902
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: 60/171,901
```

```
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: 60/171,900
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: 60/181,749
;  PRIOR FILING DATE: 2000-02-11
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 133
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
;   LENGTH: 355
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-321-807-18

  Query Match           47.5%;  Score 392.5;  DB 12;  Length 355;
  Best Local Similarity   59.1%;  Pred. No. 5.4e-35;
  Matches   75;  Conservative   25;  Mismatches   26;  Indels    1;  Gaps    1;

Qy          4 YSVAIPLGINLAAFIIVFSYGSMFTSVHQSAITATEIRNQVKKEMILAKRTFFIVTTDA 63
              ||: |||||:|| ||:|||||| :|| |: ::|:  |||:||  :|: :| ||||||:||
Db        190 YSLGIPLGVNLLAFLIVFSYIIMPCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDA 249

Qy         64 LCWIPIPVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKSMIHRFWHM 123
              :||||:|| | ||| :|||| |:|||:|! :  :|||||||||||  ||: : :  |
Db        250 ICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFKDKLQLLHK 309

Qy        124 YRQRKSM 130
                : ||||:
Db        310 H-QRKSI 315


Search completed: September 25, 2003, 15:10:13
Job time : 65 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:56:36 ; Search time 30 Seconds
(without alignments)
221.427 Million cell updates/sec

Title: US-10-049-568-2
Perfect score: 825
Sequence: 1 AQITSVAIPLGINLAAPIII.....HMLWHSSGGKCGHCRRCHLS 157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of this result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	147.5	17.9	792	3	US-07-741-453A-56	Sequence 56, Appl
2	140.5	17.0	696	3	US-07-757-342D-4	Sequence 4, Appli
3	139.5	16.9	692	3	US-07-757-342D-6	Sequence 6, Appli
4	137.5	16.6	795	3	US-07-741-453A-55	Sequence 55, Appl
5	136	16.5	370	1	US-08-118-270-18	Sequence 18, Appl
6	136	16.5	370	5	PCT-US93-08528-18	Sequence 18, Appl
7	134	16.2	700	3	US-07-757-342D-3	Sequence 3, Appli
8	131	15.9	611	3	US-07-757-342D-8	Sequence 8, Appli
9	131	15.9	636	3	US-07-757-342D-7	Sequence 7, Appli
10	131	15.9	674	3	US-07-757-342D-10	Sequence 10, Appl
11	131	15.9	699	3	US-07-757-342D-2	Sequence 2, Appli
12	125.5	15.2	420	4	US-08-795-876-33	Sequence 33, Appl
13	125.5	15.2	423	4	US-08-795-876-38	Sequence 38, Appl
14	125.5	15.2	436	4	US-08-795-876-2	Sequence 2, Appli
15	125.5	15.2	695	1	US-08-487-886-2	Sequence 2, Appli
16	125.5	15.2	695	3	US-08-482-855-2	Sequence 2, Appli
17	125.5	15.2	695	4	US-08-474-986-2	GENERAL INFORMA
18	124	15.0	338	3	US-08-988-876-8	Sequence 8, Appli
19	124	15.0	338	3	US-09-303-524A-2	Sequence 2, Appli
20	123.5	15.0	325	2	US-08-467-948A-29	Sequence 29, Appl
21	123.5	15.0	325	3	US-08-467-947A-29	Sequence 29, Appl
22	122	14.8	764	3	US-07-741-453A-54	Sequence 54, Appl
23	120.5	14.6	764	3	US-07-741-453A-60	Sequence 60, Appl
24	119.5	14.5	332	1	US-08-293-563-5	Sequence 5, Appli
25	119.5	14.5	378	3	US-09-082-088-2	Sequence 2, Appli
26	119.5	14.5	378	4	US-09-546-117-2	Sequence 2, Appli
27	119.5	14.5	378	4	US-09-169-205D-22	Sequence 22, Appl
28	119.5	14.5	378	4	US-09-731-030A-18	Sequence 18, Appl
29	119.5	14.5	517	2	US-08-467-568-10	Sequence 10, Appl
30	119.5	14.5	517	2	US-09-030-582-10	Sequence 10, Appl
31	118.5	14.3	515	1	US-08-194-338-3	Sequence 3, Appli
32	118	14.3	764	3	US-07-741-453A-61	Sequence 61, Appl
33	117.5	14.2	336	1	US-08-118-270-54	Sequence 54, Appl
34	117.5	14.2	336	5	PCT-US93-08528-54	Sequence 54, Appl
35	117.5	14.2	420	3	US-09-255-368-6	Sequence 6, Appli
36	117.5	14.2	513	2	US-08-406-855A-21	Sequence 21, Appl
37	117.5	14.2	513	2	US-09-206-899-21	Sequence 21, Appl
38	117.5	14.2	515	1	US-08-444-734A-7	Sequence 7, Appli
39	117.5	14.2	515	1	US-08-722-001-25	Sequence 25, Appl
40	117.5	14.2	515	2	US-08-406-855A-22	Sequence 22, Appl
41	117.5	14.2	515	3	US-09-206-899-22	Sequence 22, Appl
42	117.5	14.2	515	3	US-09-032-742-5	Sequence 5, Appli
43	117.5	14.2	515	4	US-09-688-415-9	Sequence 9, Appli
44	117.5	14.2	515	4	US-09-688-415-10	Sequence 10, Appl
45	117.5	14.2	764	3	US-07-757-342D-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-07-741-453A-56
; Sequence 56, Application US/07741453A
; Patent No. 6228597
;  GENERAL INFORMATION:
;     APPLICANT:  PARMENTIER, MARC
;     APPLICANT:  LIBERT, FREDERIC
;     APPLICANT:  DUMONT, JACQUES
;     APPLICANT:  VASSART, GILBERT
;     TITLE OF INVENTION:  POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
;     TITLE OF INVENTION:  ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
;     TITLE OF INVENTION:  AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
;     NUMBER OF SEQUENCES:  62
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  CUSHMAN DARBY & CUSHMAN
;       STREET:  1615 L STREET, N.W.
;       CITY:  WASHINGTON, D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/741,453A
;       FILING DATE:  19911015
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  KOKULIS, PAUL N.
;       REGISTRATION NUMBER:  16773
;       REFERENCE/DOCKET NUMBER:  91913/1107/US/ST
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 861-3000
;       TELEFAX:  (202) 822-0944
;       TELEX:  6714627 CUSH
;  INFORMATION FOR SEQ ID NO:  56:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  792 amino acids
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-07-741-453A-56

  Query Match          17.9%;  Score 147.5;  DB 3;  Length 792;
  Best Local Similarity   25.9%;  Pred. No. 1.1e-07;
  Matches   42;  Conservative   27;  Mismatches   60;  Indels   33;  Gaps    4;

Qy          1 AQITSVAIPL-----GINLAAPIIIVFSYGSMPYSVHQSAITATEIRNQVKKEMILAKRP 55
```

```
           :|:| : | :    |:|: |||||  |  ::::|    : ||       |:  :||:
Db     596 SQVYILTILIPIGHCGLNVVAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKM 649

Qy      56 FFIVFTDALCWIPIFVAKPLSLLQVE-------IPGTITSWVVIGYSAINSALNPILYTL 108
             ::|||  |  ||    |:|   :          |  | :  ::: :  :|| || || :
Db     650 AVLIFTDFTCMAPIGHCGPISFFAISAALKVPLITVTNSKVLLVLFYPVNSCANPFLYAI 709

Qy     109 TTRPFKEMIHRFWHNYRQRKSMDSKGIRKHMLHHSSGGKCGH 150
            |: |:                   |  |     |  |  | | |
Db     710 FTKAFRR---------------DPIGHCGFFLLLSKSGCCKH 736
```

```
RESULT 2
US-07-757-342D-4
; Sequence 4, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;  INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 696 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-07-757-342D-4

  Query Match             17.0%;  Score 140.5;  DB 3;  Length 696;
  Best Local Similarity   27.2%;  Pred. No. 5.6e-07;
  Matches   41;  Conservative   26;  Mismatches   55;  Indels   29;  Gaps    4;

Qy       1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
           :|:| : | : :|: |||||  |  ::::|    : ||       |:  :||:   ::|
Db     524 SQVYILTILI-LNVVAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIF 576

Qy      61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHR 119
           ||  |  || | |   :|       |  | :  ::: :  :||  || || : |: |:
Db     577 TDFTCMAPISFFAISAALKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFR----- 631

Qy     120 FWHNYRQRKSMDSKGIRKHMLHHSSGGKCGH 150
                           |   |  |  | | |
Db     632 ----------------RDFFLLLSKSGCCKH 646
```

```
RESULT 3
US-07-757-342D-6
; Sequence 6, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;  INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 692 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-07-757-342D-6

  Query Match             16.9%;  Score 139.5;  DB 3;  Length 692;
  Best Local Similarity   27.3%;  Pred. No. 7.1e-07;
  Matches   47;  Conservative   30;  Mismatches   60;  Indels   35;  Gaps    6;

Qy       1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
           :|:| :|: | :|: ||::|   | :: :|    | ::         :  :|||  ::|
Db     526 SQLYVMAL-LVLNVLAFVVICGCYTHIYLTVRNPTIVSS------SSDTKIAKRMATLIF 578

Qy      61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
           || ||  || | |   ||       |  :     ::: :  |||  || || : |: |:
Db     579 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 638

Qy     116 MIHRF----------------WHNYRQRKSMDSKGIRKH------MLHHSS 144
           :: :|                ||:  |||  |   |          |:|||
Db     639 LLSKFGCYEMQAQIYRTETSSATHNFHARKSHCSSAPRVTNSYVLVPLNHSS 690
```

```
RESULT 4
US-07-741-453A-55
; Sequence 55, Application US/07741453A
; Patent No. 6228597
;  GENERAL INFORMATION:
;    APPLICANT:  PARMENTIER, MARC
;    APPLICANT:  LIBERT, FREDERIC
;    APPLICANT:  DUMONT, JACQUES
;    APPLICANT:  VASSART, GILBERT
;    TITLE OF INVENTION:  POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
;    TITLE OF INVENTION:  ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
;    TITLE OF INVENTION:  AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
```

```
Qy          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMLAKRFFFIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMLAKRFFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTRPFKEMIHRF 120

Qy        121 WENYRQRKSMDSKG 134
              |:||||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 6
US-09-764-877-2009
; Sequence 2009, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2009
;   LENGTH: 176
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (133)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-2009

  Query Match             76.5%;  Score 632;  DB 10;  Length 176;
  Best Local Similarity   94.0%;  Pred. No. 9e-62;
  Matches  126;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMLAKRFFFIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIFLGINLAAFIIVFSYGSMFYSVHQSAITATEIRNQVKKEMLAKRFFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||    ||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYTLTRPFKEMIHRF 120

Qy        121 WENYRQRKSMDSKG 134
              |:||||||||||| |
Db        121 WYNYRQRKSMDSXG 134


RESULT 7
US-09-989-442-105
; Sequence 105, Application US/09989442
; Publication No. US20030013649A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ08
;  CURRENT APPLICATION NUMBER: US/09/989,442
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/225,757
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/226,868
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/216,647
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,267
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/216,880
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,270
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/251,869
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/235,834
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/234,274
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/234,223
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/228,924
;  PRIOR FILING DATE: 2000-08-30
;  PRIOR APPLICATION NUMBER: 60/224,518
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,369
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/224,519
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,964
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/241,809
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/249,299
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/236,327
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/241,785
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/244,617
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/225,268
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,368
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/251,856
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/251,868
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/229,344
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/234,997
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: 60/229,343
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,345
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,287
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: 60/229,513
;  PRIOR FILING DATE: 2000-09-05
;  PRIOR APPLICATION NUMBER: 60/231,413
;  PRIOR FILING DATE: 2000-09-08
;  PRIOR APPLICATION NUMBER: 60/229,509
;  PRIOR FILING DATE: 2000-09-05
```

; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14

Query Match 76.5%; Score 632; DB 11; Length 176;
Best Local Similarity 94.0%; Pred. No. 9e-62;
Matches 126; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 AQIYSVAIPLGINLAAFIIIVFSYGSMFYSVHQSAITAIEIRNQVKKEMILAKRFFFIVF 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AQIYSVAIPLGINLAAFIIIVFSYGSMFYSVHQSAITAIEIRNQVKKEMILAKRFFFIVF 60

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGVSAINSALNPILYTLTTRPFKEMIHRF 120
              |||||||||| | ||||||||||||||||||  ||||||||||||||||||||||||||||
Db         61 TDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFLLPINSALNPILYTLTTRPFKEMIHRF 120

Qy        121 WHNYRQRKSMDSKG 134
              |:||||||||| |
Db        121 WYMYRQRKSMDSXG 134
```

RESULT 8
US-09-764-886-52
; Sequence 52, Application US/09764886
; Publication No. US20030139327A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PZ%02
; CURRENT APPLICATION NUMBER: US/09/764,886
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```
RESULT 9
US-07-757-342D-7
; Sequence 7, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;  INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 636 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-07-757-342D-7

  Query Match           15.9%;  Score 131;  DB 3;  Length 636;
  Best Local Similarity  29.6%;  Pred. No. 5.4e-06;
  Matches   34;  Conservative   25;  Mismatches   48;  Indels    8;  Gaps    3;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:|: : | : :|: || ||   | ::::|    : ||       |:  :||:  :::|
Db        461 SQVYILTILI-LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIF 513

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFK 114
              || |  || | |   :    |  | :  ::: :  |||  || || : |: |:
Db        514 TDFTCMAPISFFAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 568


RESULT 10
US-07-757-342D-10
; Sequence 10, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;  INFORMATION FOR SEQ ID NO: 10:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 674 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-07-757-342D-10

  Query Match           15.9%;  Score 131;  DB 3;  Length 674;
  Best Local Similarity  29.6%;  Pred. No. 5.8e-06;
  Matches   34;  Conservative   25;  Mismatches   48;  Indels    8;  Gaps    3;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:|: : | : :|: || ||   | ::::|    : ||       |:  :||:  :::|
Db        499 SQVYILTILI-LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIF 551

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTRPFK 114
              || |  || | |   :    |  | :  ::: :  |||  || || : |: |:
Db        552 TDFTCMAPISFFAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 606


RESULT 11
US-07-757-342D-2
; Sequence 2, Application US/07757342D
; Patent No. 6218509
;  GENERAL INFORMATION:
;       APPLICANT: IGARASHI, Masao
;                  MINEGISHI, Takashi
;                  NAKAMURA, Kazuto
;       TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
;                       CUSHMAN
;            STREET: 130 Water Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: US
;            ZIP: 02109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/07/757,342D
;            FILING DATE: 10-Sep-1991
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BUCKLEY, Linda M.
;            REGISTRATION NUMBER: 31003
```

```
;            REFERENCE/DOCKET NUMBER: 41226
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)523-3400
;            TELEFAX: (617)523-6440
;            TELEX: 200291 STRE UR
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 699 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-07-757-342D-2

  Query Match           15.9%;  Score 131;  DB 3;  Length 699;
  Best Local Similarity   29.6%;  Pred. No. 6e-05;
  Matches   34;  Conservative   25;  Mismatches   48;  Indels    8;  Gaps    3;

Qy          1 AQIYSVAIPLGINLAAFIIIVFSYGSMFYSVHQSAITAYEIRNQVKKEMILAKRFFFIVF 60
              :|:| : | : :|: || ||   |  ::::|    : ||      |:  :||:   ::|
Db        524 SQVYILTILI-LNVVAFFICACYIKIYFAVRNPELKAT------NKDTKIAKKMALLIF 576

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLITRPFK 114
               || |  || | |   :     |  | :  ::: :  |||  || || : |: |:
Db        577 TDFTCMAPISFPAISAAFKVPLITVFNSKVLLVLFYPINSCANPFLYAIFTKTFQ 631

RESULT 12
US-08-795-876-33
; Sequence 33, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT: Gershengorn, Marvin C.
;    APPLICANT: Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  420 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-795-876-33

  Query Match           15.2%;  Score 125.5;  DB 4;  Length 420;
  Best Local Similarity   26.1%;  Pred. No. 1.3e-05;
  Matches   43;  Conservative   30;  Mismatches   59;  Indels   33;  Gaps    6;

Qy          1 AQIYSVAIPLGINLAAFIIIVFSYGSMFYSVHQSAITAYEIRNQVKKEMILAKRFFFIVF 60
              :|:| : | : :|: || ||   |  ::::|    : ||      |:  :||:   ::|
Db        255 SQLVVMSL-LVLNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTRIAKRMAMLIF 307

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLITRPFK----- 114
               || |  || | |   :     |  | :  ::: :  |||  || || : |: |:
Db        308 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFHPINSCANPFLYAIFTKNFRRDFPI 367

Qy         115 --------EMIHRFWHNYRQRKSMDSKGIRKHMLHSSSGGKCGHC 151
                       ||  :    ||  |              |::  : |||
Db         365 LLSKCGCYEMQAQI---YRTETSSTV---------HNTHPRNGHC 397

RESULT 13
US-08-795-876-38
; Sequence 38, Application US/08795876
; Patent No. 6403305
;  GENERAL INFORMATION:
;    APPLICANT:  Gershengorn, Marvin C.
;    APPLICANT:  Geras-Raaka, Elizabeth
;    APPLICANT:  Nussenzveig, Daniel R.
;    TITLE OF INVENTION:  STRATEGY TO CLONE DRUGS FOR G PROTEIN
;    TITLE OF INVENTION:  COUPLED RECEPTORS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON, HARGRAVE, DEVANS & DOYLE LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/795,876
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BRAMAN, SUSAN J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/1280
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-263-1636
;      TELEFAX:  716-263-1600
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  423 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-795-876-38

  Query Match           15.2%;  Score 125.5;  DB 4;  Length 423;
  Best Local Similarity   26.1%;  Pred. No. 1.3e-05;
  Matches   43;  Conservative   30;  Mismatches   59;  Indels   33;  Gaps    6;

Qy          1 AQIYSVAIPLGINLAAFIIIVFSYGSMFYSVHQSAITAYEIRNQVKKEMILAKRFFFIVF 60
              :|:| : | : :|: || ||   |  ::::|    : ||      |:  :||:   ::|
Db        255 SQLVVMSL-LVLNVLAFVVICGCYIHIYLTVRNPNIVSS------SSDTRIAKRMAMLIF 307

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLITRPFK----- 114
               || |  || | |   :     |  | :  ::: :  |||  || || : |: |:
Db        308 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFHPINSCANPFLYAIFTKNFRRDFPI 367

Qy         115 --------EMIHRFWHNYRQRKSMDSKGIRKHMLHSSSGGKCGHC 151
                       ||  :    ||  |              |::  : |||
Db         368 LLSKCGCYEMQAQI---YRTETSSTV---------HNTHPRNGHC 400
```

```
Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| : | : :|: || ||   |  ::::|     : ||       |:  :||:   ::|
Db        524 SQVYILTILI-LNVVAFFIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIF 576

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
              ||  |  || | |    :       |  | :  ::: :  |||  || || : |: |:
Db        577 TDFTCMAPISFFAISAAFKVPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFQ 631
```

Search completed: September 25, 2003, 15:01:39
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:55:26 ; Search time 36 Seconds
(without alignments)
1125.396 Million cell updates/sec

Title: US-10-049-568-2
Perfect score: 826
Sequence: 1 AQIYSVAIFLGINLAAFIII.....HMLHHSSGGXCGHCRRCHLS 157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	224.5	27.2	334	5	Q9VYG0	Q9vyg0 drosophila
2	212.5	25.7	765	5	Q9VBP0	Q9vbp0 drosophila
3	154	18.6	693	13	Q9DGC5	Q9dgc5 oreochromis
4	148	17.9	1012	5	Q95YI6	Q95yi6 asterina pe
5	148	17.9	1280	5	Q95Y17	Q95y17 asterina pe
6	143.5	17.4	696	13	Q98T85	Q98t85 ictalurus p
7	139.5	16.9	688	11	Q64183	Q64183 rattus sp.
8	138	16.7	307	13	Q90XC7	Q90xc7 salmo salar
9	134	16.2	814	13	Q91949	Q91949 oncorhynchu
10	133.5	16.2	424	13	Q9DB63	Q9db63 rana catesb
11	132	16.0	701	6	Q8SPS8	Q8sps8 bos taurus
12	132	16.0	793	13	Q91948	Q91948 oncorhynchu
13	131	15.9	779	13	Q91BM7	Q91bm7 morone saxa
14	130.5	15.8	778	13	Q98TF4	Q98tf4 oreochromis
15	130	15.7	1050	5	Q9BN18	Q9bn18 drosophila
16	130	15.7	1300	5	Q9NKD6	Q9nkd6 drosophila
17	130	15.7	1360	5	Q9NDI1	Q9ndi1 drosophila
18	128	15.5	658	13	Q9PVN9	Q9pvn9 oncorhynchu
19	127	15.4	662	13	Q98T84	Q98t84 ictalurus p
20	125.5	15.2	410	4	Q16225	Q16225 homo sapien
21	125	15.1	662	13	Q9PW16	Q9pw16 clarias gar
22	124	15.0	338	4	Q81YT7	Q81yt7 homo sapien
23	124	15.0	696	13	Q9DGF5	Q9dgf5 cynops pyrr
24	123.5	15.0	695	11	Q8R428	Q8r428 cavia porce
25	123	14.9	335	6	O46387	O46387 mustela vis
26	123	14.9	710	13	Q8AXM5	Q8axm5 clarias gar
27	122.5	14.8	364	5	Q81M7	Q81m7 asterina mi
28	122	14.8	293	6	Q95MT7	Q95mt7 canis famil
29	122	14.8	763	6	Q9BGN4	Q9bgn4 felis silve
30	121	14.6	601	13	O42500	O42500 meleagris g
31	121	14.6	764	11	Q9D697	Q9d697 mus musculu
32	119.5	14.5	384	13	Q9PUQ8	Q9puq8 fugu rubrip
33	119.5	14.5	398	4	Q9H228	Q9h228 homo sapien
34	119.5	14.5	459	11	Q8CGI5	Q8cgi5 mus musculu
35	119.5	14.5	515	11	Q9DBL0	Q9dbl0 mus musculu
36	119.5	14.5	724	13	Q9PVP0	Q9pvp0 oncorhynchu
37	119	14.4	739	6	Q9BG55	Q9bg55 sus scrofa
38	119	14.4	764	6	Q8SPP9	Q8spp9 sus scrofa
39	119	14.4	764	6	Q9BG56	Q9bg56 sus scrofa
40	117.5	14.2	418	13	Q90WJ2	Q90wj2 xenopus lae
41	117.5	14.2	470	4	O43200	O43200 homo sapien
42	117.5	14.2	533	4	Q9P1V4	Q9p1v4 homo sapien
43	117.5	14.2	764	6	Q8HY54	Q8hy54 cercopithec
44	117.5	14.2	764	6	Q8HY53	Q8hy53 macaca mula
45	116	14.0	296	11	Q8BK41	Q8bk41 mus musculu

ALIGNMENTS

RESULT 1
Q9VYG0
ID Q9VYG0 PRELIMINARY; PRT; 334 AA.
AC Q9VYG0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG4187 protein.
GN CG4187.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003491; AAF48237.1; -.
DR   FlyBase; FBgn0030458; CG4187.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001901; SecE.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01067; SECE_SEC61G; 1.
SQ   SEQUENCE   334 AA;  37020 MW;  1662F2D467534061 CRC64;

  Query Match             27.2%;  Score 224.5;  DB 5;  Length 334;
  Best Local Similarity   32.3%;  Pred. No. 5.4e-15;
  Matches   54;  Conservative   27;  Mismatches   67;  Indels   19;  Gaps    4;

Qy          4 YSVAIFLGINLAAFIIIVFSYGSMFYSVHQS--AITATEIRNQVKKEMILAKRFFFIVFT 61
              ||  :|: :|  : | |:|||  |  ::  |    : :|        :| ::| ||  || |
Db        166 YSALLFILVNTLSLIFILFSYIRMLQAIRDSGGGMRST----HSGRENVVATRFAIIVTT 221

Qy         62 DALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRFW 121
               |  ||:|| | |  :|    ||    : :|: :     :||||||:||||||  ||: : |:
Db        222 DCACWLPIIVVKLAALSGCEISPDLYAWLAVLVLPVNSALNPVLYTLTTAAFKQQLRRYC 281

Qy        122 HNY-----------RQRKSMDSKGIRKHMLHHSS--GGKCGHCRRCH 155
              |               | ::    |:   : |      ||  |  |  |
Db        282 HTLPSCSLVNNETRSQTQTAYESGLSVSLAHLGGGVGGGSGRKRMSH 328


RESULT 2
Q9VBP0
ID   Q9VBP0      PRELIMINARY;      PRT;   765 AA.
AC   Q9VBP0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG31096-PA.
GN   CG31096.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003753; AAF56490.2; -.
DR   FlyBase; FBgn0039354; Lgr3.
SQ   SEQUENCE   765 AA;  87214 MW;  48FC480611F8BFD6 CRC64;

  Query Match             25.7%;  Score 212.5;  DB 5;  Length 765;
  Best Local Similarity   35.9%;  Pred. No. 2.2e-13;
  Matches   47;  Conservative   24;  Mismatches   55;  Indels    5;  Gaps    3;

Qy          3 IYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTD 62
               :||  :|||:||    ::|    | ::  |: ::  :|| :        :   | ||||||  ||
Db        605 LYSAFVFLGVNLLLLVMIAMLYTALLISIWRTR-SATPL---TLLDCEFAVRFFFIVLTD 660
```

```
Qy          63 ALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIH-RFW 121
               |||:|| | |            |    | :|:|:     :|||:||:||| ||   ::   |   | |
Db         661 FLCWVPIIVMKIWVFFNYNISDDIYAWLVVFVLPLNSAVNPLLYTFTTPKYRNQIFLRGW 720

Qy         122 HNYRQRKSMDS 132
                     ||  ::
Db         721 KKITSRKRAEA 731
```

RESULT 3
Q9DGC5

```
ID   Q9DGC5      PRELIMINARY;      PRT;   693 AA.
AC   Q9DGC5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Gonadotropin receptor II.
GN   TGTH-RII.
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oba Y., Hirai T., Yoshiura Y., Yao Z., Nagahama Y.;
RT   "Tilapia gonadotropin receptor II.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB041763; BAB16107.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   693 AA;  77856 MW;  C3BBEFD8ECFC8988 CRC64;

  Query Match             18.6%;  Score 154;  DB 13;  Length 693;
  Best Local Similarity   34.5%;  Pred. No. 2.2e-07;
  Matches   40;  Conservative   24;  Mismatches   44;  Indels    8;  Gaps    4;

Qy           1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
               ||:| ||: : :|: ||::: : |   :: |||      :      |    |     :|||    ::|
Db         524 AQVYVVAVLI-LNVVAFLVVCYCYICIYLSVHNPEHSTR--RGDTK----IAKRMAVLIF 576

Qy          61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE 115
               || ||   || | |    :|     |  : :   ::| :  |||   || |||: || |::
Db         577 TDFLCMAPISFFAISAALRMPLITVSHSKILLILFYPINSLCNPFLYTIFTTRAFRK 632
```

RESULT 4
Q95YI6

```
ID   Q95YI6      PRELIMINARY;      PRT;  1012 AA.
AC   Q95YI6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycoprotein hormone receptor.
GN   APGPHR'.
OS   Asterina pectinifera (Starfish).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC   Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX   NCBI_TaxID=7594;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
RT   "cDNA cloning and functional analysis of a novel member of the
RT   glycoprotein hormone receptor family from a starfish Asterina
RT   pectinifera.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB061862; BAB68209.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 13.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 5.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS50506; LRR_TYPICAL; 3.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   1012 AA;  112623 MW;  52A70E7A88C46E0A CRC64;

  Query Match             17.9%;  Score 148;  DB 5;  Length 1012;
  Best Local Similarity   32.9%;  Pred. No. 1.3e-06;
  Matches   51;  Conservative   27;  Mismatches   51;  Indels   26;  Gaps    9;

Qy           1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQS-AITATEIRNQVKKEMILAKRFFFIV 59
               |::| |    | :|: ||:||:  | |:: ::  | |      : |          :|:|   :|
Db         719 AKVY-VGSILILNILAFVIIMACYASIYLAIQGSHAWNCNDSR--------VARRMSLLV 769

Qy          60 FTDALCWIPI-----FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
               ||| ||| ||          |   | |: ::      :| :|:      :||  || |||: |: ||
Db         770 FTDFACWAPIAFFSLTAAFGLRLISLDGAKVLTIFVL----PLNSCANPFLYTILTKQFK 825

Qy         115 ---EMIHRFWHN--YRQRKSMDSKGIRKHMLHHSS 144
                    : | |   |  :||| || |: |   :  | |
Db         826 KDCKTIMRSLSNRVFRQR-SM-SRSITLSLGRHPS 858
```

RESULT 5
Q95YI7

```
ID   Q95YI7      PRELIMINARY;      PRT;  1280 AA.
AC   Q95YI7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycoprotein hormone receptor.
GN   APGPHR.
OS   Asterina pectinifera (Starfish).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC   Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
OX   NCBI_TaxID=7594;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
RT   "cDNA cloning and functional analysis of a novel member of the
RT   glycoprotein hormone receptor family from a starfish Asterina
RT   pectinifera.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB061861; BAB68208.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 13.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   SMART; SM00369; LRR_TYP; 5.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS50506; LRR_TYPICAL; 3.
```

```
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   1280 AA;  141700 MW;  0AB0ECCC0DD880BA CRC64;

  Query Match             17.9%;  Score 148;  DB 5;  Length 1280;
  Best Local Similarity   32.9%;  Pred. No. 1.7e-06;
  Matches   51;  Conservative   27;  Mismatches   51;  Indels   26;  Gaps    9;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQS-AITATEIRNQVKKEMILAKRFFFIV 59
                |::| |   | :|: ||:||: | |:: ::  | |     : |         :|:|  :|
Db        719 AKVY-VGSILILNILAFVIIMACYASIYLAIQGSHAWNCNDSR--------VARRMSLLV 769

Qy         60 FTDALCWIPI-----FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
              |||  || ||       | | |: ::    :| :|:      :||  || |||: |: ||
Db        770 FTDFACWAPIAFFSLTAAFGLRLISLDGAKVLTIFVL----PLNSCANPFLYTILTKQFK 825

Qy        115 ---EMIHRFWHN--YRQRKSMDSKGIRKHMLHHSS 144
                 : | |    |  :||| || |: |    :  | |
Db        826 KDCKTIMRSLSNRVFRQR-SM-SRSITLSLGRHPS 858


RESULT 6
Q98T85
ID   Q98T85      PRELIMINARY;      PRT;   696 AA.
AC   Q98T85;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Luteinizing hormone receptor precursor.
OS   Ictalurus punctatus (Channel catfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC   Ictaluridae; Ictalurus.
OX   NCBI_TaxID=7998;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21141975; PubMed=11207219;
RA   Sampath Kumar R., Ijiri S., Trant J.M.;
RT   "Molecular Biology of Channel Catfish Gonadotropin Receptors: 1.
RT   Cloning of a Functional Luteinizing Hormone Receptor and Preovulatory
RT   Induction of Gene Expression.";
RL   Biol. Reprod. 64:1010-1018(2001).
CC   -!- SUBCELLULAR LOCATION: INIEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF285181; AAK16066.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR000307; Ribosomal_S16.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00732; RIBOSOMAL_S16; 1.
KW   G-protein coupled receptor; Receptor; Signal; Transmembrane.
FT   SIGNAL        1     23       POTENTIAL.
SQ   SEQUENCE   696 AA;  77822 MW;  4A1FB19CDE070AB6 CRC64;

  Query Match             17.4%;  Score 143.5;  DB 13;  Length 696;
  Best Local Similarity   29.5%;  Pred. No. 2.6e-06;
  Matches   49;  Conservative   30;  Mismatches   66;  Indels   21;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :| | ||: | ||:|||:::  ||  :: ||    :         :       :|||   ::|
Db        527 SQGYVVAV-LVINVAAFLMVCSSYAGIYLSVRNPNVPTRHGHTR------MAKRMAVLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |    :|     |  : :  ::| :  |||  || |||: || |:
Db        580 TDFLCMAPISFFAISAALHMPLISVSQSKILLILFYPINSLCNPFLYTIFTRAFRRDMCL 639

Qy        116 MIHRFWHNYRQRKSMDSKGI-------RKHMLH--HSSGGKCGHCR 152
              :: |    : | :  |:|:         ||  :  ||   :  | :
Db        640 LLSRCGCCHAQAEFYRSQGLLGFTLAPRKKRVRKPHSKNFRAYHVK 685


RESULT 7
Q64183
ID   Q64183      PRELIMINARY;      PRT;   688 AA.
AC   Q64183;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Follicle-stimulating hormone receptor.
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92149579; PubMed=1738373;
RA   Heckert L.L., Daley I.J., Griswold M.D.;
RT   "Structural organization of the follicle-stimulating hormone receptor
RT   gene.";
RL   Mol. Endocrinol. 6:70-80(1992).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; S81198; AAB21415.2; -.
DR   EMBL; S81117; AAB21415.2; JOINED.
DR   EMBL; S81119; AAB21415.2; JOINED.
DR   EMBL; S81121; AAB21415.2; JOINED.
DR   EMBL; S81171; AAB21415.2; JOINED.
DR   EMBL; S81174; AAB21415.2; JOINED.
DR   EMBL; S81178; AAB21415.2; JOINED.
DR   EMBL; S81183; AAB21415.2; JOINED.
DR   EMBL; S81185; AAB21415.2; JOINED.
DR   EMBL; S81194; AAB21415.2; JOINED.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 3.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   688 AA;  77341 MW;  441F0D9E7D01DF18 CRC64;

  Query Match             16.9%;  Score 139.5;  DB 11;  Length 688;
  Best Local Similarity   27.3%;  Pred. No. 6.8e-06;
  Matches   47;  Conservative   30;  Mismatches   60;  Indels   35;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| :|: | :|: ||::|    |  :: :|     | ::            :  :|||   ::|
Db        522 SQLYVMAL-LVLNVLAFVVICGCYTHIYLTVRNPTIVSS------SSDTKIAKRMATLIF 574

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGIITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |    ||      |  :     ::: :  |||  || || : |: |:
Db        575 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 634

Qy        116 MIHRF----------------WHNYRQRKSMDSKGIRKH------MLHHSS 144
              :: :|                ||:  |||  |   |      |:|||
Db        635 LLSKFGCYEMQAQIYRTETSSATHNFHARKSHCSSAPRVTNSYVLVPLNHSS 686


RESULT 8
Q90XC7
ID   Q90XC7      PRELIMINARY;      PRT;   307 AA.
AC   Q90XC7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative thyrotropin receptor TSHR (Fragment).
OS   Salmo salar (Atlantic salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX   NCBI_TaxID=8030;
```

```
FT   DISULFID     79     91       BY SIMILARITY.
FT   DISULFID     86    104       BY SIMILARITY.
FT   DISULFID     98    113       BY SIMILARITY.
FT   DISULFID    118    131       BY SIMILARITY.
FT   DISULFID    138    153       BY SIMILARITY.
FT   DISULFID    158    170       BY SIMILARITY.
FT   DISULFID    165    183       BY SIMILARITY.
FT   DISULFID    177    194       BY SIMILARITY.
FT   DISULFID    202    220       BY SIMILARITY.
FT   DISULFID    214    230       BY SIMILARITY.
FT   DISULFID    233    245       BY SIMILARITY.
FT   DISULFID    240    258       BY SIMILARITY.
FT   DISULFID    252    267       BY SIMILARITY.
FT   DISULFID    274    291       BY SIMILARITY.
FT   DISULFID    282    304       BY SIMILARITY.
FT   DISULFID    298    316       BY SIMILARITY.
FT   DISULFID    322    339       BY SIMILARITY.
FT   DISULFID    334    352       BY SIMILARITY.
FT   DISULFID    346    361       BY SIMILARITY.
FT   DISULFID    367    379       BY SIMILARITY.
FT   DISULFID    374    392       BY SIMILARITY.
FT   DISULFID    386    401       BY SIMILARITY.
FT   DISULFID    406    418       BY SIMILARITY.
FT   DISULFID    413    431       BY SIMILARITY.
FT   DISULFID    425    440       BY SIMILARITY.
FT   DISULFID    446    458       BY SIMILARITY.
FT   DISULFID    453    474       BY SIMILARITY.
FT   DISULFID    465    483       BY SIMILARITY.
FT   DISULFID    488    500       BY SIMILARITY.
FT   DISULFID    495    513       BY SIMILARITY.
FT   DISULFID    507    523       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    166    166       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    269    269       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    318    318       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    482    482       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    502    502       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    571    571       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    618    618       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    624    624       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    685    685       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1115 AA;  125865 MW;  E9BB01297ECE356C CRC64;

  Query Match             25.1%;  Score 207.5;  DB 1;  Length 1115;
  Best Local Similarity   33.1%;  Pred. No. 1.5e-11;
  Matches   51;  Conservative   25;  Mismatches   45;  Indels   33;  Gaps    5;

Qy          4 YSVAIFLGINLAAFIIIVFSYGSMFYSV---HQSAITATEIRNQVKKEMILAKRFFFIVF 60
                ||| :|: :|| :|::|  ||  || ||     :||:   | :|     :  :|:|   ||
Db        938 YSVGVFILLNLLSFVLIASSYLWMF-SVAKKTRSAVRTAESKN----DNAMARRMTLIVM 992

Qy         61 TDALCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPF------- 113
                ||   ||:|| |    :||             : :|: :     :||| ||::|||:| ||
Db        993 TDFCCWVPIIVLGFVSLAGARADDQVYAWIAVFVLPLNSATNPVIYTLSTAPFLGNVRKR 1052

Qy        114 -----KEMIHRF-------------WHNYRQRKS 129
                    |  || |              |:| ::||
Db       1053 ANRFRKSFIHSFTGDTKHSYVDDGTTHSYCEKKS 1086


RESULT 5
FSHR_SHEEP
ID   FSHR_SHEEP     STANDARD;      PRT;   695 AA.
AC   P35379; Q28573; Q28574; Q9TSI9;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE   receptor).
GN   FSHR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM FSH-R1).
RC   TISSUE=Testis;
RX   MEDLINE=93351750; PubMed=8394255;
RA   Yarney T.A., Sairam M.R., Khan H., Ravindranath N., Payne S.,
RA   Seidah N.G.;
RT   "Molecular cloning and expression of the ovine testicular follicle
RT   stimulating hormone receptor.";
RL   Mol. Cell. Endocrinol. 93:219-226(1993).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS FSH-R4 AND FSH-R3).
RC   STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
RX   MEDLINE=93176195; PubMed=8439338;
RA   Khan H., Yarney T.A., Sairam M.R.;
RT   "Cloning of alternately spliced mRNA transcripts coding for variants
RT   of ovine testicular follitropin receptor lacking the G protein
RT   coupling domains.";
RL   Biochem. Biophys. Res. Commun. 190:888-894(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
RC   STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Testis;
RX   MEDLINE=98031015; PubMed=9364440;
RA   Yarney T.A., Jiang L., Khan H., MacDonald E.A., Laird D.W.,
RA   Sairam M.R.;
RT   "Molecular cloning, structure, and expression of a testicular
RT   follitropin receptor with selective alteration in the carboxy terminus
RT   that affects signaling function.";
RL   Mol. Reprod. Dev. 48:458-470(1997).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
RC   STRAIN=Dorset-Leicester-Suffolk 1; TISSUE=Ovary;
RX   MEDLINE=20391225; PubMed=10527886;
RA   Babu P.S., Jiang L., Sairam A.M., Touyz R.M., Sairam M.R.;
RT   "Structural features and expression of an alternatively spliced growth
RT   factor type I receptor for follitropin signaling in the developing
RT   ovary.";
RL   Mol. Cell Biol. Res. Commun. 2:21-27(1999).
RN   [5]
RP   SEQUENCE OF 1-51 FROM N.A.
RX   MEDLINE=98031017; PubMed=9364442;
RA   Sairam M.R., Subbarayan V.S.R.;
RT   "Characterization of the 5' flanking region and potential control
RT   elements of the ovine follitropin receptor gene.";
RL   Mol. Reprod. Dev. 48:480-487(1997).
CC   -!- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC       of isoform FSH-R1 is mediated by G proteins which activate
CC       adenylate cyclase. Isoforms FSH-R2 and FSH-R3 also bind FSH, but
CC       this does not result in activation of adenylate cyclase. Isoform
CC       FSH-R3 may be involved in calcium signaling.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
CC       (isoforms FSH-R1 and FSH-R2); Cell surface (isoform FSH-R3).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=FSH-R1;
CC         IsoId=P35379-1; Sequence=Displayed;
CC       Name=FSH-R2;
CC         IsoId=P35379-2; Sequence=VSP_001959, VSP_001960;
CC       Name=FSH-R3;
CC         IsoId=P35379-3; Sequence=VSP_001957, VSP_001958;
CC       Name=FSH-R4;
CC         IsoId=P35379-4; Sequence=VSP_001955, VSP_001956;
CC   -!- TISSUE SPECIFICITY: Isoform FSH-R3 is expressed in ovary and
CC       testis, but not in kidney.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L07302; AAA31525.1; -.
DR   EMBL; L12766; AAA31523.1; -.
DR   EMBL; L12767; AAA31524.1; -.
DR   EMBL; L36115; AAK70667.1; -.
DR   EMBL; AJ131735; CAA10495.1; -.
DR   EMBL; AF090438; AAC61749.1; -.
DR   PIR; JC1493; JC1493.
DR   HSSP; P23945; 1XUN.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 5.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    695       FOLLICLE STIMULATING HORMONE RECEPTOR.
FT   DOMAIN       18    366       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    367    387       1 (POTENTIAL).
FT   DOMAIN      388    398       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    399    421       2 (POTENTIAL).
FT   DOMAIN      422    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    444    465       3 (POTENTIAL).
FT   DOMAIN      466    485       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    486    508       4 (POTENTIAL).
FT   DOMAIN      509    528       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    529    550       5 (POTENTIAL).
FT   DOMAIN      551    573       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    574    597       6 (POTENTIAL).
FT   DOMAIN      598    608       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    609    630       7 (POTENTIAL).
FT   DOMAIN      631    695       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       44     68       LRR 1.
FT   REPEAT       69     93       LRR 2.
FT   REPEAT      119    143       LRR 3.
FT   REPEAT      170    192       LRR 4.
FT   REPEAT      193    216       LRR 5.
FT   REPEAT      218    240       LRR 6.
FT   DISULFID    442    517       BY SIMILARITY.
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    126    133       LISNTGIK -> FKRWRNRI (in isoform FSH-R4).
FT                                /FTId=VSP_001955.
FT   VARSPLIC    135    695       Missing (in isoform FSH-R4).
FT                                /FTId=VSP_001956.
FT   VARSPLIC    224    259       DISRTRIRSLPSYGLENLKKLRAKSTYHLKKLPSLE -> S
FT                                PLLHCWAHLQSFFFVVCGQREHISEFGLKSKQHPN (in
FT                                isoform FSH-R3).
FT                                /FTId=VSP_001957.
FT   VARSPLIC    260    695       Missing (in isoform FSH-R3).
FT                                /FTId=VSP_001958.
FT   VARSPLIC    643    670       KFGCYEVQAQTYRSETSFTAHNFHPRNG -> LHCCTVGLI
FT                                CNHFSSLFVARGNIFLNLD (in isoform FSH-R2).
FT                                /FTId=VSP_001959.
FT   VARSPLIC    671    695       Missing (in isoform FSH-R2).
FT                                /FTId=VSP_001960.
SQ   SEQUENCE   695 AA;  78237 MW;  FBF75D89D88C0D4B CRC64;

  Query Match             17.4%;  Score 143.5;  DB 1;  Length 695;
  Best Local Similarity   27.2%;  Pred. No. 7.7e-06;
  Matches   44;  Conservative   32;  Mismatches   59;  Indels   27;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::: | :|: ||::|   |  :: :|    ||::         :  :|||   ::|
Db        527 SQLYVMSL-LVLNVLAFVVICGCYTHIYLTVRNPNITSS------SSDTKIAKRMAMLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
                || ||  || | |   ||     |  : :  ::: :  |||  || || : || |:
Db        580 TDFLCMAPISFFAISASLKVPLITVSKSKILLVLFYPINSCANPFLYAIFTRNFRRDFFI 639

Qy        116 MIHRFW------HNYRQRKSMDSKGIRKHMLHHSSGGKCGHC 151
                :: :|               ||    |  :            |:   :  |||
Db        640 LLSKFGCYEVQAQTYRSETSFTA---------HNFHPRNGHC 672


RESULT 6
GLHR_ANTEL
ID   GLHR_ANTEL     STANDARD;      PRT;   925 AA.
AC   P35409;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable glycoprotein hormone G-protein coupled receptor precursor.
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RX   MEDLINE=94107299; PubMed=8280121;
RA   Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT   "Molecular cloning of a novel, putative G protein-coupled receptor
RT   from sea anemones structurally related to members of the FSH, TSH,
RT   LH/CG receptor family from mammals.";
RL   Biochem. Biophys. Res. Commun. 197:1062-1069(1993).
CC   -!- FUNCTION: PROBABLE RECEPTOR FOR A GLYCOPROTEIN HORMONE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P35409-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=P35409-2; Sequence=VSP_001983;
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z28332; CAA82186.1; -.
DR   PIR; JC2033; JC2033.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 6.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Alternative splicing; Repeat.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    925       PROBABLE GLYCOPROTEIN HORMONE G-PROTEIN
FT                                COUPLED RECEPTOR.
FT   DOMAIN       28    529       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    530    551       1 (POTENTIAL).
FT   DOMAIN      552    561       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    562    584       2 (POTENTIAL).
FT   DOMAIN      585    606       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    607    628       3 (POTENTIAL).
FT   DOMAIN      629    651       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM    652    673       4 (POTENTIAL).
FT   DOMAIN      674    691       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    692    712       5 (POTENTIAL).
FT   DOMAIN      713    739       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    740    763       6 (POTENTIAL).
FT   DOMAIN      764    774       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    775    795       7 (POTENTIAL).
FT   DOMAIN      796    925       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      333    461       5 X APPROXIMATE TANDEM REPEATS.
FT   REPEAT      333    349       1 (INCOMPLETE).
FT   REPEAT      350    384       2.
FT   REPEAT      385    419       3.
FT   REPEAT      420    453       4.
FT   REPEAT      454    461       5 (INCOMPLETE).
FT   CARBOHYD     61     61       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    152    152       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    212    212       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    435    435       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    495    495       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    235    925       Missing (in isoform Short).
FT                                /FTId=VSP_001983.
SQ   SEQUENCE   925 AA;  100059 MW;  D03A256368452FBD CRC64;

  Query Match             17.3%;  Score 143;  DB 1;  Length 925;
  Best Local Similarity   33.3%;  Pred. No. 1.2e-05;
  Matches   37;  Conservative   21;  Mismatches   47;  Indels    6;  Gaps    3;

Qy          6 VAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKK-EMILAKRFFFIVFTDAL 64
              ||  | :| |:|| ::: |  | | |    ::   ::     |: :  :|||    :|||| |
Db        694 VAFLLLVNGASFISVMYLYSRMLYVV----VSGGDMEGAPKRNDSKVAKRMAILVFTDML 749

Qy         65 CWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
              || || |      :  |   :   | :  ::: :  |||  || ||    |: ||
Db        750 CWAPIAFFGLLAAFGQTLLTVTQSKILLVFFFPINSICNPFLYAFFTKAFK 800


RESULT 7
FSHR_MACFA
ID   FSHR_MACFA     STANDARD;      PRT;   695 AA.
AC   P32212;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE   receptor).
GN   FSHR.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=94071854; PubMed=7504463;
RA   Gromoll J., Dankbar B., Sharma R.S., Nieschlag E.;
RT   "Molecular cloning of the testicular follicle stimulating hormone
RT   receptor of the non human primate Macaca fascicularis and
RT   identification of multiple transcripts in the testis.";
RL   Biochem. Biophys. Res. Commun. 196:1066-1072(1993).
CC   -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC       OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC       ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X74454; CAA52463.1; -.
DR   PIR; JN0898; JN0898.
DR   HSSP; P23945; 1XUN.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 3.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    695       FOLLICLE STIMULATING HORMONE RECEPTOR.
FT   DOMAIN       18    366       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    367    387       1 (POTENTIAL).
FT   DOMAIN      388    398       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    399    421       2 (POTENTIAL).
FT   DOMAIN      422    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    444    465       3 (POTENTIAL).
FT   DOMAIN      466    485       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    486    508       4 (POTENTIAL).
FT   DOMAIN      509    528       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    529    550       5 (POTENTIAL).
FT   DOMAIN      551    573       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    574    597       6 (POTENTIAL).
FT   DOMAIN      598    608       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    609    630       7 (POTENTIAL).
FT   DOMAIN      631    695       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       44     68       LRR 1.
FT   REPEAT       69     93       LRR 2.
FT   REPEAT      119    143       LRR 3.
FT   REPEAT      170    192       LRR 4.
FT   REPEAT      193    216       LRR 5.
FT   REPEAT      218    240       LRR 6.
FT   DISULFID    442    517       BY SIMILARITY.
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    318    318       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   695 AA;  78343 MW;  0D60A233729B5250 CRC64;

  Query Match             17.1%;  Score 141.5;  DB 1;  Length 695;
  Best Local Similarity   26.5%;  Pred. No. 1.2e-05;
  Matches   44;  Conservative   32;  Mismatches   63;  Indels   27;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| ::: | :|: ||::|    |  :: :|     | ::            :  :|||    ::|
Db        527 SQLYVMSL-LVLNVLAFVVICGCYTHIYLTVRNPNIVSS------SSDTRIAKRMAMLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |    ||      |  :     ::: :  |||  || ||  : |: |:
Db        580 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 639

Qy        116 MIHRFW------HNYRQRKSMDSKGIRKHMLHHSSGGKCGHCRRCH 155
              :: :|            ||    |  :              |:|   : |||    |
Db        640 LLSKFGCYEMQAQIYRTETSSTA---------HNSHPRNGHCSSAH 676


RESULT 8
LSHR_PIG
ID   LSHR_PIG       STANDARD;      PRT;   696 AA.
AC   P16582;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
```

```
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D).
RX   MEDLINE=89332517; PubMed=2502844;
RA   Loosfelt H., Misrahi M., Atger M., Salesse R., Thi M.T.V.H.-L.,
RA   Jolivet A., Guiochon-Mantel A., Sar S., Jallal B., Garnier J.,
RA   Milgrom E.;
RT   "Cloning and sequencing of porcine LH-hCG receptor cDNA: variants
RT   lacking transmembrane domain.";
RL   Science 245:525-528(1989).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE. THE
CC       ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC       ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=A;
CC         IsoId=P16582-1; Sequence=Displayed;
CC       Name=B;
CC         IsoId=P16582-2; Sequence=VSP_001963, VSP_001964;
CC       Name=C;
CC         IsoId=P16582-3; Sequence=VSP_001965, VSP_001966;
CC       Name=D;
CC         IsoId=P16582-4; Sequence=VSP_001967;
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to licensei@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M29525; AAA31062.1; -.
DR   EMBL; M29526; AAA31063.1; -.
DR   EMBL; M29527; AAA31064.1; -.
DR   EMBL; M29528; AAA31065.1; -.
DR   PIR; A41344; A41344.
DR   PIR; B41344; B41344.
DR   PIR; C41344; C41344.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT   SIGNAL      1     27       POTENTIAL.
FT   CHAIN      28    696       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                              RECEPTOR.
FT   DOMAIN     28    358       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  359    386       1 (POTENTIAL).
FT   DOMAIN    387    395       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM  396    418       2 (POTENTIAL).
FT   DOMAIN    419    439       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  440    462       3 (POTENTIAL).
FT   DOMAIN    463    482       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM  483    505       4 (POTENTIAL).
FT   DOMAIN    506    525       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  526    547       5 (POTENTIAL).
FT   DOMAIN    548    570       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM  571    594       6 (POTENTIAL).
FT   DOMAIN    595    605       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  606    626       7 (POTENTIAL).
FT   DOMAIN    627    696       CYTOPLASMIC (POTENTIAL).
FT   REPEAT     48     71       LRR 1.
FT   REPEAT    122    147       LRR 2.
FT   REPEAT    149    171       LRR 3.
FT   REPEAT    172    196       LRR 4.
FT   REPEAT    198    220       LRR 5.
FT   REPEAT    221    244       LRR 6.
FT   DISULFID  439    514       BY SIMILARITY.
FT   CARBOHYD   93     93       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  174    174       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  291    291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  299    299       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  313    313       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC  317    329       YSAIPAESELSDW -> LLHGALPATHCLS (in
FT                              isoform B).
FT                              /FTId=VSP_001963.
FT   VARSPLIC  330    696       Missing (in isoform B).
FT                              /FTId=VSP_001964.
FT   VARSPLIC  317    331       YSAIPAESELSDWDY -> SKSRADGYQQRHKDC (in
FT                              isoform C).
FT                              /FTId=VSP_001965.
FT   VARSPLIC  332    696       Missing (in isoform C).
FT                              /FTId=VSP_001966.
FT   VARSPLIC  317    628       Missing (in isoform D).
FT                              /FTId=VSP_001967.
SQ   SEQUENCE   696 AA;  78092 MW;  593DEF1C25F982FE CRC64;

  Query Match           17.0%;  Score 140.5;  DB 1;  Length 696;
  Best Local Similarity  27.2%;  Pred. No. 1.5e-05;
  Matches   41;  Conservative   26;  Mismatches   55;  Indels   29;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMPYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| : | : :|: |||||   |  ::::|    : ||       |: :|:|:  ::|
Db        524 SQVYILTLLI-LNVVAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIF 576

Qy         61 TDALCWIPI-PVAKPLSLLQVEIPCTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHR 119
              || |  || | |    :|    |  | :  ::: :  :||  || || : |: |:
Db        577 TDFTCMAFISFFAISAALKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFR----- 631

Qy        120 FWHNYRQRXSMDSKGIRKHMLEHSSGGKCGH 150
                                |  |  |  | | |
Db        632 ----------------RDFFLLLSKSGCCKH 646

RESULT 9
FSHR_RAT
ID   FSHR_RAT       STANDARD;      PRT;   692 AA.
AC   P20395;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE   receptor).
GN   FSHR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Sertoli cells;
RX   MEDLINE=91125358; PubMed=2126341;
RA   Sprengel R., Braun T., Nikolics K., Segaloff D.L., Seeburg P.H.;
RT   "The testicular receptor for follicle stimulating hormone: structure
RT   and functional expression of cloned cDNA.";
RL   Mol. Endocrinol. 4:525-530(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 25, 2003, 14:54:41 ; Search time 11 Seconds
(without alignments)
671.199 Million cell updates/sec

Title: US-10-049-568-2
Perfect score: 826
Sequence: 1 AQIYSVAIFLGINLAAFIII.....RNLHHSSGGKCGHCRRCHLS 157

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	638	77.2	757	1	LGR7_HUMAN	Q9HBX9 homo sapien
2	409	49.5	737	1	LGR8_MOUSE	Q91zz5 mus musculu
3	392.5	47.5	754	1	LGR8_HUMAN	Q8wxd0 homo sapien
4	207.5	25.1	1115	1	GPCR_LYMST	P46023 lymnaea sta
5	143.5	17.4	695	1	FSHR_SHEEP	P35379 ovis aries
6	143	17.3	925	1	GLHR_ANTEL	P35409 anthopleura
7	141.5	17.1	695	1	FSHR_MACFA	P32212 macaca fasc
8	140.5	17.0	696	1	LSHR_PIG	P16582 sus scrofa
9	139.5	16.9	692	1	FSHR_RAT	P20395 rattus norv
10	139.5	16.9	695	1	FSHR_BOVIN	P35376 bos taurus
11	139	16.8	676	1	LSHR_CALJA	O02721 callithrix
12	136.5	16.5	695	1	FSHR_PIG	P49059 sus scrofa
13	134	16.2	700	1	LSHR_MOUSE	P30730 mus musculu
14	134	16.2	700	1	LSHR_RAT	P16235 rattus norv
15	132.5	16.0	692	1	FSHR_MOUSE	P35375 mus musculu
16	132	16.0	701	1	LSHR_BOVIN	Q28005 bos taurus
17	131.5	15.9	694	1	FSHR_HORSE	P47799 equus cabal
18	131	15.9	693	1	FSHR_CHICK	P79763 gallus gall
19	131	15.9	699	1	LSHR_HUMAN	P22888 homo sapien
20	129	15.6	687	1	FSHR_EQUAS	Q95179 equus asinu
21	128.5	15.6	319	1	AA3R_MOUSE	Q61618 mus musculu
22	128	15.5	320	1	AA3R_RAT	P28647 rattus norv
23	127	15.4	764	1	TSHR_SHEEP	P56495 ovis aries
24	125.5	15.2	695	1	FSHR_HUMAN	P23945 homo sapien
25	124	15.0	338	1	P2YX_HUMAN	Q15391 homo sapien
26	124	15.0	763	1	TSHR_BOVIN	Q27987 bos taurus
27	124	15.0	764	1	TSHR_RAT	P21463 rattus norv
28	123	14.9	340	1	AA2B_CHICK	O13076 gallus gall
29	123	14.9	366	1	LSHR_CHICK	Q90674 gallus gall
30	122	14.8	764	1	TSHR_CANFA	P14763 canis famil
31	121	14.6	764	1	TSHR_MOUSE	P47750 mus musculu
32	119.5	14.5	332	1	AA2B_RAT	P29276 rattus norv
33	119.5	14.5	378	1	EDG3_HUMAN	O99500 homo sapien
34	119.5	14.5	417	1	A1AB_CANFA	P11615 canis famil
35	119	14.4	519	1	A1AB_HUMAN	P35368 homo sapien
36	118.5	14.3	317	1	AA3R_SHEEP	P93542 ovis aries
37	118.5	14.3	338	1	P2YX_MOUSE	Q9esq6 mus musculu
38	118	14.3	332	1	AA2B_MOUSE	Q60614 mus musculu
39	117.5	14.2	514	1	A1AB_MOUSE	P97717 mus musculu
40	117.5	14.2	515	1	A1AB_MESAU	P18841 mesocricetu
41	117.5	14.2	515	1	A1AB_RAT	P15823 rattus norv
42	117.5	14.2	522	1	NPF2_HUMAN	Q9y5x5 homo sapien
43	117	14.2	326	1	AA1R_BOVIN	P28190 bos taurus
44	117	14.2	326	1	AA1R_CAVPO	P47745 cavia porce
45	117	14.2	332	1	AA2B_HUMAN	P29275 homo sapien

ALIGNMENTS

```
RESULT 1
LGR7_HUMAN
ID   LGR7_HUMAN      STANDARD;      PRT;   757 AA.
AC   Q9HBX9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Relaxin receptor 1 (Leucine-rich repeat-containing G protein-coupled
DE   receptor 7).
GN   LGR7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF ASP-637.
RX   MEDLINE=20388592; PubMed=10935549;
RA   Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA   van der Spek P.J., van Duin M., Hsueh A.J.W.;
RT   "The three subfamilies of leucine-rich repeat-containing G protein-
RT   coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT   signaling mechanism for LGR7.";
RL   Mol. Endocrinol. 14:1257-1271(2000).
RN   [2]
RP   CHARACTERIZATION.
RX   MEDLINE=21408393; PubMed=11517286;
RA   Bartsch O., Bartlick B., Ivell R.;
RT   "Relaxin signalling links tyrosine phosphorylation to
RT   phosphodiesterase and adenylyl cyclase activity.";
RL   Mol. Hum. Reprod. 7:799-809(2001).
CC   -!- FUNCTION: Receptor for relaxin. The activity of this receptor is
CC       mediated by G proteins leading to stimulation of adenylate cyclase
CC       and an increase of cAMP. Binding of the ligand may also activate a
CC       tyrosine kinase pathway that inhibits the activity of a
CC       phosphodiesterase that degrades cAMP.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9HBX9-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9HBX9-2; Sequence=VSP_001984;
CC   -!- TISSUE SPECIFICITY: Expressed in the brain, kidney, testis,
CC       placenta, uterus, ovary, adrenal, prostate, skin and heart. Not
CC       detected in spleen.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC   -!- SIMILARITY: Contains 10 leucine-rich (LRR) repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; AF190500; AAG17167.1; -.
DR   MIM; 606654; -.
DR   HSSP; P01130; 1LDR.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_receptor_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 1.
DR   Pfam; PF00560; LRR; 8.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00192; LDLa; 1.
DR   SMART; SM00369; LRR_TYP; 2.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat; Alternative splicing.
FT   DOMAIN      1    409       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  410    430       1 (POTENTIAL).
FT   DOMAIN    431    443       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM  444    464       2 (POTENTIAL).
FT   DOMAIN    465    486       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  487    507       3 (POTENTIAL).
FT   DOMAIN    508    527       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM  528    548       4 (POTENTIAL).
FT   DOMAIN    549    577       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  578    598       5 (POTENTIAL).
FT   DOMAIN    599    629       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM  630    650       6 (POTENTIAL).
FT   DOMAIN    651    651       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM  652    672       7 (POTENTIAL).
FT   DOMAIN    673    757       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     26     63       LDL-RECEPTOR CLASS A.
FT   REPEAT    125    148       LRR 1.
FT   REPEAT    150    172       LRR 2.
FT   REPEAT    173    196       LRR 3.
FT   REPEAT    198    220       LRR 4.
FT   REPEAT    221    244       LRR 5.
FT   REPEAT    246    269       LRR 6.
FT   REPEAT    270    293       LRR 7.
FT   REPEAT    294    317       LRR 8.
FT   REPEAT    319    341       LRR 9.
FT   REPEAT    343    365       LRR 10.
FT   DISULFID  485    563       BY SIMILARITY.
FT   CARBOHYD   36     36       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  264    264       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  272    272       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  325    325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  368    368       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   63     96       GDNNGWSMQPDKTFASTYKMTSQVPFEAETPECL -> V
FT                              (in isoform 2).
FT                              /FTId=VSP_001984.
FT   MUTAGEN   637    637       D->Y: LEADS TO CONSTITUTIVE INCREASE OF
FT                              BASAL CAMP.
SQ   SEQUENCE   757 AA;  86992 MW;  8079E8DBF3AJEF21 CRC64;

  Query Match          77.2%;  Score 638;  DB 1;  Length 757;
  Best Local Similarity   94.8%;  Pred. No. 1.6e-51;
  Matches  127;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         1 AQIYSVAIPLGINLAAPIIVFSYIGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       576 AQIYSVAIPLGINLAAPIIVFSYIGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 635

Qy        61 TDALCWIPIFVAKFLSLLQVEIPQTITSWVVIGYSAINSALNPILYTLTNTRPFKEMIHRF 120
             |||||||||| | ||||||||||||||||||  ||||||||||||||||||||||||||||
Db       636 TDALCWIPIFVVKFLSLLQVEIPQTITSWVVIPILPINSALNPILYTLTNTRPFKEMIHRF 695

Qy       121 WHNYRQRKSMDSKG 134
             ||||||||||||||
Db       696 WTNYRQRKSMDSKG 709


RESULT 2
LGR8_MOUSE
ID   LGR8_MOUSE      STANDARD;      PRT;   737 AA.
AC   Q91Z25;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Relaxin receptor 2 (Leucine-rich repeat-containing G protein-coupled
DE   receptor 8) (G protein-coupled receptor affecting testicular descent)
DE   (G protein-coupled receptor 106).
GN   LGR8 OR GREAT OR GPR106.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=21250990; PubMed=11353515;
RA   Overbeek P.A., Gorlov I.P., Sutherland R.W., Houston J.B.,
RA   Harrison W.R., Boettger-Tong H.L., Bishop C.E., Agoulnik A.I.;
RT   "A transgenic insertion causing cryptorchidism in mice.";
RL   Genesis 30:26-35(2001).
CC   -!- FUNCTION: Receptor for relaxin. The activity of this receptor is
CC       mediated by G proteins leading to stimulation of adenylate cyclase
CC       and an increase of cAMP. May also be a receptor for Leydig
CC       insulin-like peptide (INSL3) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in embryonic and adult gonads of
CC       males and females, as well in male gubernaculum. Expressed also
CC       in brain. Not detected in kidney, spleen and heart.
CC   -!- DISEASE: Defects in LGR8 seems to be a cause of impaired
CC       testicular descent (known as cryptorchidism).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC   -!- SIMILARITY: Contains 10 leucine-rich (LRR) repeats.
CC   -!- CAUTION: By homology with the human sequence, it is uncertain
CC       whether Met-1 is the initiator.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF346501; AAL08943.1; -.
DR   MGD; MGI:2153463; Gpr106.
DR   GO; GO:0016500; F:protein-hormone receptor activity; IMP.
DR   GO; GO:0007189; P:G-protein signaling, adenylate cyclase acti. . .; IMP.
DR   GO; GO:0008584; P:male gonad development; IMP.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_receptor_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 1.
DR   Pfam; PF00560; LRR; 7.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00192; LDLa; 1.
DR   SMART; SM00369; LRR_TYP; 2.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 1.
```

```
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat.
FT   DOMAIN        1    399       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    400    420       1 (POTENTIAL).
FT   DOMAIN      421    438       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    439    459       2 (POTENTIAL).
FT   DOMAIN      460    490       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    491    511       3 (POTENTIAL).
FT   DOMAIN      512    520       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    521    541       4 (POTENTIAL).
FT   DOMAIN      542    575       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    576    596       5 (POTENTIAL).
FT   DOMAIN      597    622       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    623    643       6 (POTENTIAL).
FT   DOMAIN      644    644       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    645    665       7 (POTENTIAL).
FT   DOMAIN      666    737       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       27     64       LDL-RECEPTOR CLASS A.
FT   REPEAT      119    142       LRR 1.
FT   REPEAT      144    166       LRR 2.
FT   REPEAT      167    190       LRR 3.
FT   REPEAT      192    214       LRR 4.
FT   REPEAT      215    238       LRR 5.
FT   REPEAT      239    262       LRR 6.
FT   REPEAT      263    286       LRR 7.
FT   REPEAT      287    310       LRR 8.
FT   REPEAT      312    334       LRR 9.
FT   REPEAT      335    358       LRR 10.
FT   DISULFID    478    556       BY SIMILARITY.
FT   CARBOHYD     37     37       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    121    121       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    257    257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    318    318       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    361    361       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   737 AA;  82943 MW;  AF60F635EA1ACE49 CRC64;

  Query Match             49.5%;  Score 409;  DB 1;  Length 737;
  Best Local Similarity   58.9%;  Pred. No. 2.3e-30;
  Matches   73;  Conservative   27;  Mismatches   24;  Indels    0;  Gaps    0;

Qy          4 YSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDA 63
              ||: ||||:|| ||::|| || :|| |:|::|:   |:|: : ||: :| |||||||:||
Db        572 YSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAVANRFFFIVFSDA 631

Qy         64 LCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRFWHN 123
              :||||:|| | |||||||||||||||:|:  : :|||||||||||||  ||: : :  | 
Db        632 ICWIPVFVVKILSLLQVEIPGTITSWIVVFFLPVNSALNPILYTLTTSFFKDKLKQLLHK 691

Qy        124 YRQR 127
              :|::
Db        692 HRRK 695


RESULT 3
LGR8_HUMAN
ID   LGR8_HUMAN     STANDARD;      PRT;   754 AA.
AC   Q8WXD0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Relaxin receptor 2 (Leucine-rich repeat-containing G protein-coupled
DE   receptor 8) (G protein-coupled receptor affecting testicular descent).
GN   LGR8 OR GREAT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF ASP-647.
RX   MEDLINE=21669315; PubMed=11809971;
RA   Hsu S.Y., Nakabayashi K., Nishi S., Kumagai J., Kudo M.,
RA   Sherwood O.D., Hsueh A.J.W.;
RT   "Activation of orphan receptors by the hormone relaxin.";
RL   Science 295:671-674(2002).
RN   [2]
RP   SEQUENCE FROM N.A., VARIANT CRYPTORCHIDISM PRO-222, AND VARIANT
RP   VAL-604.
RX   MEDLINE=22206409; PubMed=12217959;
RA   Gorlov I.P., Kamat A., Bogatcheva N.V., Jones E., Lamb D.J.,
RA   Truong A., Bishop C.E., McElreavey K., Agoulnik A.I.;
RT   "Mutations of the GREAT gene cause cryptorchidism.";
RL   Hum. Mol. Genet. 11:2309-2318(2002).
CC   -!- FUNCTION: Receptor for relaxin. The activity of this receptor is
CC       mediated by G proteins leading to stimulation of adenylate cyclase
CC       and an increase of cAMP. May also be a receptor for Leydig
CC       insulin-like peptide (INSL3).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed mainly in the brain kidney, muscle,
CC       testis, thyroid,uterus, peripheral blood cells and bone marrow.
CC   -!- DISEASE: Defects in LGR8 are the a cause of cryptorchidism [MIM:
CC       219050]. This failure of testicular descent is one of the most
CC       frequent congenital malformations, affecting 1-3% of newborn boys.
CC       Cryptorchidism is associated with increased risk of infertility
CC       and testicular cancer.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC   -!- SIMILARITY: Contains 10 leucine-rich (LRR) repeats.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-18 is the initiator.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF403384; AAL69324.2; -.
DR   EMBL; AF453828; AAL73946.1; -.
DR   MIM; 606655; -.
DR   MIM; 219050; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_receptor_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 1.
DR   Pfam; PF00560; LRR; 7.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00192; LDLa; 1.
DR   SMART; SM00369; LRR_TYP; 3.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 1.
DR   PROSITE; PS50068; LDLRA_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat; Polymorphism; Disease mutation.
FT   DOMAIN        1    416       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    417    437       1 (POTENTIAL).
FT   DOMAIN      438    455       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    456    476       2 (POTENTIAL).
FT   DOMAIN      477    507       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    508    528       3 (POTENTIAL).
FT   DOMAIN      529    537       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    538    558       4 (POTENTIAL).
FT   DOMAIN      559    592       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    593    613       5 (POTENTIAL).
FT   DOMAIN      614    639       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    640    660       6 (POTENTIAL).
FT   DOMAIN      661    673       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    674    694       7 (POTENTIAL).
FT   DOMAIN      695    754       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       44     81       LDL-RECEPTOR CLASS A.
```

```
FT   REPEAT        136     159       LRR 1.
FT   REPEAT        161     183       LRR 2.
FT   REPEAT        184     207       LRR 3.
FT   REPEAT        209     231       LRR 4.
FT   REPEAT        233     255       LRR 5.
FT   REPEAT        256     279       LRR 6.
FT   REPEAT        281     303       LRR 7.
FT   REPEAT        304     327       LRR 8.
FT   REPEAT        329     351       LRR 9.
FT   REPEAT        352     375       LRR 10.
FT   DISULFID      495     573       BY SIMILARITY.
FT   CARBOHYD       54      54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      138     138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      274     274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      335     335       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD      378     378       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT       222     222       T -> P (in cryptorchidism; functionally
FT                                   inactive).
FT                                   /FTId=VAR_015386.
FT   VARIANT       604     604       I -> V.
FT                                   /FTId=VAR_015387.
FT   MUTAGEN       647     647       D->Y: LEADS TO CONSTITUTIVE INCREASE OF
FT                                   BASAL CAMP.
SQ   SEQUENCE   754 AA;  86452 MW;  2088ECD204C6A6C5 CRC64;

  Query Match             47.5%;  Score 392.5;  DB 1;  Length 754;
  Best Local Similarity   59.1%;  Pred. No. 8e-29;
  Matches  75;  Conservative   25;  Mismatches   26;  Indels    1;  Gaps    1;

Qy          4 YSVAIFLGINLAAFIIIVFSYYSSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVFTDA 63
              ||: ||||:|| ||:|||||| :|| |: :::|:  ||:||   :|: :| ||||||:||
Db        589 YSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDA 648

Qy         64 LCWIPIFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHRFWHN 123
              :||||:|| | ||| :|||| :|||:|| : :|||||||||||||  ||: : :  |
Db        649 ICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYTLTTNFFKDLKQLLSK 708

Qy         124 YRQRKSM 130
                : ||||:
Db         709 H-QRKSI 714


RESULT 4
GPCR_LYMST
ID   GPCR_LYMST     STANDARD;      PRT;  1115 AA.
AC   P46023;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   G-protein coupled receptor GRL101 precursor.
OS   Lymnaea stagnalis (Great pond snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC   Lymnaeoidea; Lymnaeidae; Lymnaea.
OX   NCBI_TaxID=6523;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=CNS;
RX   MEDLINE=94255418; PubMed=8197140;
RA   Tensen C.P., van Kesteren E.R., Planta R.J., Cox K.J.A., Burke J.F.,
RA   van Heerikhuizen H., Vreugdenhil E.;
RT   "A G protein-coupled receptor with low density lipoprotein-binding
RT   motifs suggests a role for lipoproteins in G-linked signal
RT   transduction.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:4816-4820(1994).
CC   -!- FUNCTION: MIGHT DIRECTLY TRANSDUCE SIGNALS CARRIED BY LARGE
CC       EXTRACELLULAR (LIPO)PROTEIN(COMPLEXE)S INTO NEURONAL EVENTS.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN A SMALL NUMBER OF
CC       NEURONS WITHIN THE CENTRAL NERVOUS SYSTEM AND TO A LESSER EXTENT
CC       IN THE HEART.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC   -!- SIMILARITY: Contains 12 LDL-receptor class A domains.
CC   -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z23104; CAA80651.1; -.
DR   PIR; S40241; S40241.
DR   HSSP; Q07954; 1CR8.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR002172; LDL_receptor_A.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00057; ldl_recept_a; 11.
DR   Pfam; PF00560; LRR; 6.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   SMART; SM00192; LDLa; 12.
DR   SMART; SM00369; LRR_TYP; 1.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS01209; LDLRA_1; 6.
DR   PROSITE; PS50068; LDLRA_2; 11.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW   Leucine-rich repeat; Signal.
FT   SIGNAL        1      24       POTENTIAL.
FT   CHAIN        25    1115       G-PROTEIN COUPLED RECEPTOR GRL101.
FT   DOMAIN       25     767       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    768     788       1 (POTENTIAL).
FT   DOMAIN      789     801       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    802     822       2 (POTENTIAL).
FT   DOMAIN      823     857       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    858     878       3 (POTENTIAL).
FT   DOMAIN      879     887       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    888     908       4 (POTENTIAL).
FT   DOMAIN      909     941       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    942     962       5 (POTENTIAL).
FT   DOMAIN      963     988       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    989    1009       6 (POTENTIAL).
FT   DOMAIN     1010    1017       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1018    1038       7 (POTENTIAL).
FT   DOMAIN     1039    1115       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       32     523       12 X 40 AA APPROXIMATE TANDEM REPEATS
FT                                 SIMILAR TO THE LDL-RECEPTOR CLASS A.
FT   DOMAIN       36      79       LDL-RECEPTOR CLASS A 1.
FT   DOMAIN       77     115       LDL-RECEPTOR CLASS A 2.
FT   DOMAIN      116     155       LDL-RECEPTOR CLASS A 3.
FT   DOMAIN      156     196       LDL-RECEPTOR CLASS A 4.
FT   DOMAIN      195     232       LDL-RECEPTOR CLASS A 5.
FT   DOMAIN      231     269       LDL-RECEPTOR CLASS A 6.
FT   DOMAIN      272     318       LDL-RECEPTOR CLASS A 7.
FT   DOMAIN      320     363       LDL-RECEPTOR CLASS A 8.
FT   DOMAIN      365     403       LDL-RECEPTOR CLASS A 9.
FT   DOMAIN      404     442       LDL-RECEPTOR CLASS A 10.
FT   DOMAIN      444     485       LDL-RECEPTOR CLASS A 11.
FT   DOMAIN      486     525       LDL-RECEPTOR CLASS A 12.
FT   REPEAT      557     579       LRR 1.
FT   REPEAT      582     605       LRR 2.
FT   REPEAT      607     629       LRR 3.
FT   REPEAT      630     653       LRR 4.
FT   REPEAT      655     677       LRR 5.
FT   REPEAT      678     701       LRR 6.
FT   REPEAT      703     725       LRR 7.
FT   DISULFID     38      53       BY SIMILARITY.
FT   DISULFID     46      66       BY SIMILARITY.
FT   DISULFID     60      77       BY SIMILARITY.
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Yadetie F., Male R.;
RT   "Cloning and characterization of Atlantic salmon ovarian thyrotropin
RT   receptor (TSHR) like cDNA.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF406603; AAK98600.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
FT   NON_TER      1      1
SQ   SEQUENCE   307 AA;  34289 MW;  0E3B8032DAF68470 CRC64;

  Query Match             16.7%;  Score 138;  DB 13;  Length 307;
  Best Local Similarity   28.9%;  Pred. No. 4.1e-06;
  Matches   35;  Conservative   25;  Mismatches   41;  Indels   20;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              ||:| |:: : :|: ||::|   |  ::|:||      :        |    :|||    :|
Db        115 AQVYIVSVLI-LNILAFLVICTCYMKIYYAVHNPYYCSG------SKNTNIAKRMAIFIF 167

Qy         61 TDALCWIPI-------FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPF 113
              || || ||        | :||     |  :  : ::: :  :||  || || : |: |
Db        168 TDFLCMAPISFYAMSPVVDRPL------ITVSNSKILLVLFYPLNSCANPFLYAIFTKAF 221

Qy        114 K 114
              :
Db        222 R 222


RESULT 9
Q9I949
ID   Q9I949      PRELIMINARY;      PRT;   814 AA.
AC   Q9I949;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Thyrotropin receptor A.
GN   STSH-RA.
OS   Oncorhynchus rhodurus (amago).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=41164;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20462948; PubMed=11006115;
RA   Oba Y., Hirai T., Yoshiura Y., Kobayashi T., Nagahama Y.;
RT   "Cloning, Functional Characterization, and Expression of Thyrotropin
RT   Receptors in the Thyroid of Amago Salmon (Oncorhynchus rhodurus).";
RL   Biochem. Biophys. Res. Commun. 276:258-263(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB030954; BAB07800.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   814 AA;  91091 MW;  1FC9A2AE8D6C911E CRC64;

  Query Match             16.2%;  Score 134;  DB 13;  Length 814;
  Best Local Similarity   26.8%;  Pred. No. 3e-05;
  Matches   44;  Conservative   32;  Mismatches   60;  Indels   28;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              ||:| |:: : :|: ||::|   |  :: :||      :        |:  :|||    ::|
Db        624 AQVYIVSVLI-LNILAFLVICACYIKIYCTVHNPHYRSG------SKDTNIAKRMAVLIF 676

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK----- 114
              || ||  || | |     | :  |  :  :  ::: :  :||  || || : |: |:
Db        677 TDFLCMAPISFYAMSAVLDRPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFRGDVFI 736

Qy        115 -----EMIHRFWHNYRQRKSMDSKGIRKHMLHHSSGGKCGHCRR 153
                   :  |    :|  ::: |||        | | |   ||
Db        737 LLSKVGLCQRRAQLFR-GQTVSSKG---------SSGVCHQGRR 770


RESULT 10
Q9DE63
ID   Q9DE63      PRELIMINARY;      PRT;   424 AA.
AC   Q9DE63;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GnRH receptor-3 (Type III GnRHR).
GN   BFGNRHR-3.
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=21065155; PubMed=11120886;
RA   Wang L., Bogerd J., Choi H.S., Seong J.Y., Soh J.M., Chun S.Y.,
RA   Blomenrohr M., Troskie B.E., Millar R.P., Yu W.H., McCann S.M.,
RA   Kwon H.B.;
RT   "Three distinct types of GnRH receptor characterized in the
RT   bullfrog.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:361-366(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21407809; PubMed=11517181;
RA   Wang L., Oh D.Y., Bogerd J., Choi H.S., Ahn R.S., Seong J.Y.,
RA   Kwon H.B.;
RT   "Inhibitory activity of alternative splice variants of the bullfrog
RT   GnRH receptor-3 on wild-type receptor signaling.";
RL   Endocrinology 142:4015-4025(2001).
DR   EMBL; AF144062; AAG42574.1; -.
DR   EMBL; AF224277; AAL11631.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   424 AA;  47613 MW;  E771F7403AB44AD6 CRC64;

  Query Match             16.2%;  Score 133.5;  DB 13;  Length 424;
  Best Local Similarity   22.6%;  Pred. No. 1.7e-05;
  Matches   44;  Conservative   33;  Mismatches   73;  Indels   45;  Gaps    6;

Qy          3 IYSVAIFLGINLAAFIIIVFSYGSMFYSV----HQSAITATEI------RNQVKKEMILA 52
              :|:: |  : |   :|:|| || :   :      :: ::: |:        |  :  |
Db        224 LYNMFTFCCLFLLPLLIMVFCYGRILVEISRKMKKAEVSSREVNLRRSYNNIPRARMRTF 283

Qy         53 KRFFFIVFTDALCWIPIFVA------KPLSLLQVEIPGTITSWVVIGYSAINSALNPILY 106
              |    || |  :|| | ::          |  |    ::| :: | ::  :    |: |:||:|
Db        284 KMSLVIVLTFIVCWTPYYLLGIWYWFSPEMLTSRKVPPSL-SHILFLFGLFNTCLDPIIY 342

Qy        107 TLTTRPFKEMIHRFWHNYRQRKSMDSKGIRKHM---------LHHSSG------------ 145
              | |  |: | |          | |  |:  :                 |  |:|
Db        343 GLFTIHFRREIRRVCRCATQGKDADATSLGTGSFRISTAAVPLKRSAGASGGSCKFDLEV 402

Qy        146 -------GKCGHCRR 153
                     ||| ||:|
Db        403 TGVGLHSGKCEHCKR 417
```

```
RESULT 11
Q8SPS8
ID   Q8SPS8      PRELIMINARY;      PRT;    701 AA.
AC   Q8SPS8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Luteinizing hormone receptor.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein; TISSUE=Corpus luteum;
RA   Kawate N., Tamada H., Inaba T., Sawada T.;
RT   "Molecular cloning of a full-length cDNA encoding bovine luteinizing
RT   hormone receptor and its expression in COS-7 cells.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF491303; AAM09535.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 3.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   701 AA;  78455 MW;  D70AE862EB265CCF CRC64;

  Query Match             16.0%;  Score 132;  DB 6;  Length 701;
  Best Local Similarity   29.6%;  Pred. No. 4.1e-05;
  Matches   34;  Conservative   26;  Mismatches   47;  Indels    8;  Gaps    3;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| : | : :|: |||||   |  ::::|      : ||        |:  :||:    ::|
Db        526 SQVYILTILI-LNVMAFIIICACYIKIYFAVQNPELMAT------NKDTKIAKKMAVLIF 578

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
              ||  |  || | |    :       |  |  :  ::: :   :||  || || : |: |:
Db        579 TDFTCMAPISFFAISAAFKVPLITVTNSKVLLVLFYPVNSCANPFLYAIFTKAFQ 633


RESULT 12
Q9I948
ID   Q9I948      PRELIMINARY;      PRT;    793 AA.
AC   Q9I948;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Thyrotropin receptor B.
GN   STSH-RB.
OS   Oncorhynchus rhodurus (amago).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=41164;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20462948; PubMed=11006115;
RA   Oba Y., Hirai T., Yoshiura Y., Kobayashi T., Nagahama Y.;
RT   "Cloning, Functional Characterization, and Expression of Thyrotropin
RT   Receptors in the Thyroid of Amago Salmon (Oncorhynchus rhodurus).";
RL   Biochem. Biophys. Res. Commun. 276:258-263(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB030955; BAB07801.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   793 AA;  89410 MW;  2842BF0D67ED93DC CRC64;

  Query Match             16.0%;  Score 132;  DB 13;  Length 793;
  Best Local Similarity   27.3%;  Pred. No. 4.7e-05;
  Matches   33;  Conservative   28;  Mismatches   40;  Indels   20;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              ||:| ::: : :|: ||::|   |  :: :||       :          |:  :|||    ::|
Db        598 AQVYIISVLI-LNILAFLVICTCYVKIYCAVHNPYYCSG------SKDTNIAKRMAILIF 650

Qy         61 TDALCWIPI-------FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPF 113
              || ||  ||         | :||       |  : :  ::: :· :||  || || : |: |
Db        651 TDFLCMAPISFYAMSAVVDRPL------ITVSNSKILLVLFYPLNSCANPFLYAIFTKAF 704

Qy        114 K 114
              :
Db        705 R 705


RESULT 13
Q9I8N7
ID   Q9I8N7      PRELIMINARY;      PRT;    779 AA.
AC   Q9I8N7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Thyrotropin receptor precursor.
OS   Morone saxatilis (Striped bass).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Moronidae; Morone.
OX   NCBI_TaxID=34816;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=20457109; PubMed=11000515;
RA   Kumar R.S., Ijiri S., Kight K., Swanson P., Dittman A., Alok D.,
RA   Zohar Y., Trant J.M.;
RT   "Cloning and functional expression of a thyrotropin receptor from the
RT   gonads of a vertebrate (bony fish): potential thyroid-independent role
RT   for thyrotropin in reproduction.";
RL   Mol. Cell. Endocrinol. 167:1-9(2000).
DR   EMBL; AF239761; AAF80596.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor; Signal.
FT   SIGNAL        1     22       POTENTIAL.
SQ   SEQUENCE   779 AA;  87400 MW;  8C86D121B676A9EF CRC64;

  Query Match             15.9%;  Score 131;  DB 13;  Length 779;
  Best Local Similarity   29.6%;  Pred. No. 5.8e-05;
  Matches   34;  Conservative   27;  Mismatches   46;  Indels    8;  Gaps    3;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              ||:| ::: | :|: ||::|   |  :: :||       :          |:  :|||    ::|
Db        588 AQVYILSV-LVLNILAFVVICACYFKIYCAVHNPHYRSG------SKDTNIAKRMAVLIF 640

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
              || || || | |    | :  |  : :  ::: :  :||  || || : |: |:
```

```
Db        641 TDFLCMAPISFYAMSAVLDRPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFR 695


RESULT 14
Q98TF4
ID   Q98TF4      PRELIMINARY;      PRT;   778 AA.
AC   Q98TF4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Thyrotropin receptor.
GN   TTSH-R.
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oba Y., Nagahama Y.;
RT   "tilapia thyrotropin receptor.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AB047390; BAB39132.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   778 AA;  87126 MW;  618A314295880A9E CRC64;

  Query Match             15.8%;  Score 130.5;  DB 13;  Length 778;
  Best Local Similarity   28.0%;  Pred. No. 6.5e-05;
  Matches   49;  Conservative   30;  Mismatches   67;  Indels   29;  Gaps    7;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              |||| ::: | :|: || ||   |  :: :||           :   |:  :|||   ::|
Db        587 AQIYIMSV-LVLNILAFFIICACYFKIYCTVHNPHY------HSGSKDTNIAKRMAVLIF 639

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKEMIHR 119
              || ||  || | |     | :   |  : :  ::: :  :||  || || : |: |: |
Db        640 TDFLCMAPISFYAISAVLDRPLITVSNSKILLVLFYPLNSCANPFLYAIFTKAFRGDIFI 699

Qy        120 FWHNY---RQR------KSMDSKG------IRKHMLHHSSGGKCG------HCRR 153
                         :||       ::: |||      ||:       ||  |       | :|
Db        700 LLSKVGLCQQRAQLFRGQTVSSKGSSGTSQIRRDKDKSRKGGSIGQEEVPIHLKR 754


RESULT 15
Q9BN18
ID   Q9BN18      PRELIMINARY;      PRT;  1050 AA.
AC   Q9BN18;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leucine-rich repeat-containing G protein-coupled receptor 2.
GN   RK OR CG8930.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nishi S., Hsu S.Y., Zell K., Hsueh A.J.;
RT   "Characterization of two fly LGR (leucine-rich repeat-containing G
RT   protein-coupled receptor) proteins homologous to vertebrate
RT   glycoprotein hormone receptors: constitutively activation of wild type
RT   fly LGR1 but not LGR2 in transfected mammalian cells.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF274591; AAK00808.1; -.
DR   FlyBase; FBgn0003255; rk.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 10.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00369; LRR_TYP; 4.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS50506; LRR_TYPICAL; 2.
KW   Receptor.
SQ   SEQUENCE   1050 AA;  117707 MW;  35D71260A8B4BF99 CRC64;

  Query Match             15.7%;  Score 130;  DB 5;  Length 1050;
  Best Local Similarity   30.9%;  Pred. No. 0.0001;
  Matches   38;  Conservative   21;  Mismatches   42;  Indels   22;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              | :  |    : ||  ||: ::  |  |::::  |     |          :  :||| :||
Db        704 ASLTYVISLMFINGCAFLTLMGCYLKMYWAIRGSQAWNT-------NDSRIAKRMALLVF 756

Qy         61 TDALCWIP--------IFVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRP 112
              || ||| |        ||  : :|| | :|   | :|:    :||  || || : |:
Db        757 TDFLCWSPIAFFSITAIFGLQLISLEQAKI---FTVFVL----PLNSCCNPFLYAIMTKQ 809

Qy        113 FKE 115
              ||:
Db        810 FKK 812


Search completed: September 25, 2003, 15:00:38
Job time : 38 secs
```

```
RX   MEDLINE=92149579; PubMed=1738373;
RA   Heckert L.L., Daley I.J., Griswold M.D.;
RT   "Structural organization of the follicle-stimulating hormone receptor
RT   gene.";
RL   Mol. Endocrinol. 6:70-80(1992).
CC   -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC       OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC       ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: SERTOLI CELLS AND OVARIAN GRANULOSA CELLS.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L02842; AAA41175.1; -.
DR   PIR; A34548; A34548.
DR   HSSP; P23945; 1XUN.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 3.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    692       FOLLICLE STIMULATING HORMONE RECEPTOR.
FT   DOMAIN       18    365       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    366    386       1 (POTENTIAL).
FT   DOMAIN      387    397       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    398    420       2 (POTENTIAL).
FT   DOMAIN      421    442       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    443    464       3 (POTENTIAL).
FT   DOMAIN      465    484       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    485    507       4 (POTENTIAL).
FT   DOMAIN      508    527       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    528    549       5 (POTENTIAL).
FT   DOMAIN      550    572       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    573    596       6 (POTENTIAL).
FT   DOMAIN      597    607       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    608    629       7 (POTENTIAL).
FT   DOMAIN      630    692       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       44     68       LRR 1.
FT   REPEAT       69     93       LRR 2.
FT   REPEAT      119    143       LRR 3.
FT   REPEAT      170    192       LRR 4.
FT   REPEAT      193    216       LRR 5.
FT   REPEAT      218    240       LRR 6.
FT   DISULFID    441    516       BY SIMILARITY.
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   692 AA;  77681 MW;  267EA78C7CFD8EC6 CRC64;

  Query Match             16.9%;  Score 139.5;  DB 1;  Length 692;
  Best Local Similarity   27.3%;  Pred. No. 1.8e-05;
  Matches   47;  Conservative   30;  Mismatches   60;  Indels   35;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| :|: | :|: ||::|    |  :: :|    | ::          :  :|||   ::|
Db        526 SQLYVMAL-LVLNVLAFVVICGCYTHIYLTVRNPTIVSS------SSDTKIAKRMATLIF 578

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |    ||      |  :      ::: :   |||   || || : |: |:
Db        579 TDFLCMAPISFFAISASLKVPLITVSKAKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 638

Qy        116 MIHRF-----------------WHNYRQRKSMDSKGIRKH------MLHHSS 144
              :: :|                 ||:  |||  |    |            |:|||
Db        639 LLSKFGCYEMQAQIYRTETSSATHNFHARKSHCSSAPRVTNSYVLVPLNHSS 690

RESULT 10
FSHR_BOVIN
ID   FSHR_BOVIN     STANDARD;      PRT;   695 AA.
AC   P35376;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE   receptor).
GN   FSHR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Holstein; TISSUE=Ovary, and Testis;
RX   MEDLINE=95127199; PubMed=7826612;
RA   Houde A., Lambert A., Saumande J., Silversides D.W., Lussier J.G.;
RT   "Structure of the bovine follicle-stimulating hormone receptor
RT   complementary DNA and expression in bovine tissues.";
RL   Mol. Reprod. Dev. 39:127-135(1994).
CC   -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC       OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC       ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L22319; AAC37324.1; -.
DR   PIR; I45896; I45896.
DR   HSSP; P23945; 1XUN.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 5.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    695       FOLLICLE STIMULATING HORMONE RECEPTOR.
FT   DOMAIN       18    366       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    367    387       1 (POTENTIAL).
FT   DOMAIN      388    398       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    399    421       2 (POTENTIAL).
FT   DOMAIN      422    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    444    465       3 (POTENTIAL).
FT   DOMAIN      466    485       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM    486    508       4 (POTENTIAL).
FT   DOMAIN      509    528       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    529    550       5 (POTENTIAL).
FT   DOMAIN      551    573       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    574    597       6 (POTENTIAL).
FT   DOMAIN      598    608       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    609    630       7 (POTENTIAL).
FT   DOMAIN      631    695       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       44     68       LRR 1.
FT   REPEAT       69     93       LRR 2.
FT   REPEAT      119    143       LRR 3.
FT   REPEAT      170    192       LRR 4.
FT   REPEAT      193    216       LRR 5.
FT   REPEAT      218    240       LRR 6.
FT   DISULFID    442    517       BY SIMILARITY.
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   695 AA;  78084 MW;  18F9DFEFC046380D CRC64;

  Query Match             16.9%;  Score 139.5;  DB 1;  Length 695;
  Best Local Similarity   26.5%;  Pred. No. 1.8e-05;
  Matches   43;  Conservative   33;  Mismatches   59;  Indels   27;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::: | :|: ||::|   |  :: :|    ||::       :  :||   ::|
Db        527 SQLYVMSL-LVLNVLAFVVICGCYTHIYLTVRNPNITSS------SSDTKIAKRMAMLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |    ||     |  :  :  ::: :  |||  || || : |: |:
Db        580 TDFLCMAPISFFAISASLKVPLITVSKSKILLVLFYPINSCANPFLYAIFTKNFRRDFFI 639

Qy        116 MIHRFW------HNYRQRKSMDSKGIRKHMLHHSSGGKCGHC 151
              :: :|          ||   |   :          |:    : |||
Db        640 LLSKFGCYEVQAQTYRSETSSTA---------HNFHPRNGHC 672


RESULT 11
LSHR_CALJA
ID   LSHR_CALJA      STANDARD;      PRT;   676 AA.
AC   O02721;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR.
OS   Callithrix jacchus (Common marmoset).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC   Callithrix.
OX   NCBI_TaxID=9483;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=97307697; PubMed=9165039;
RA   Zhang F.-P., Rannikko A.S., Manna P.R., Fraser H.M., Huhtaniemi I.T.;
RT   "Cloning and functional expression of the luteinizing hormone
RT   receptor complementary deoxyribonucleic acid from the marmoset monkey
RT   testis: absence of sequences encoding exon 10 in other species.";
RL   Endocrinology 138:2481-2490(1997).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U80673; AAB53698.1; -.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     29       POTENTIAL.
FT   CHAIN        30    676       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                                RECEPTOR.
FT   DOMAIN       30    340       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    341    362       1 (POTENTIAL).
FT   DOMAIN      363    372       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    373    393       2 (POTENTIAL).
FT   DOMAIN      394    416       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    417    439       3 (POTENTIAL).
FT   DOMAIN      440    459       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    460    482       4 (POTENTIAL).
FT   DOMAIN      483    502       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    503    526       5 (POTENTIAL).
FT   DOMAIN      527    547       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    548    571       6 (POTENTIAL).
FT   DOMAIN      572    582       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    583    604       7 (POTENTIAL).
FT   DOMAIN      605    676       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       52     75       LRR 1.
FT   REPEAT      126    151       LRR 2.
FT   REPEAT      153    175       LRR 3.
FT   REPEAT      176    200       LRR 4.
FT   REPEAT      225    248       LRR 5.
FT   DISULFID    416    491       BY SIMILARITY.
FT   CARBOHYD    103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    178    178       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   676 AA;  75677 MW;  FB018227641654E1 CRC64;

  Query Match             16.8%;  Score 139;  DB 1;  Length 676;
  Best Local Similarity   31.3%;  Pred. No. 2e-05;
  Matches   36;  Conservative   24;  Mismatches   47;  Indels    8;  Gaps    3;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| : | : :|: |||||  |  ::::|     : ||          |:  :||:   ::|
Db        501 SQIYILTILI-LNVVAFIIICACYIKIYFAVRNPELMAT------NKDTKIAKKMAILIF 553

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFK 114
              ||  |  || | |     :       |  | :  ::: :  |||  || || : |: |:
Db        554 TDFTCMAPISFFAISAAFKMPLITVTNSKVLLVLFYPINSCANPFLYAIFTKTFR 608


RESULT 12
FSHR_PIG
ID   FSHR_PIG       STANDARD;      PRT;   695 AA.
AC   P49059; O77514;
DT   01-FEB-1996 (Rel. 33, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE   receptor).
GN   FSHR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=96011644; PubMed=7590277;
RA   Remy J.J., Lahbib-Mansais Y., Yerle M., Bozon V., Couture L.,
RA   Pajot E., Grebert D., Salesse R.;
RT   "The porcine follitropin receptor: cDNA cloning, functional
RT   expression and chromosomal localization of the gene.";
RL   Gene 163:257-261(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F.,
RA   la Barbera A.R.;
RT   "Porcine follicle-stimulating hormone receptor.";
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC       OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC       ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L31966; AAA86933.1; -.
DR   EMBL; AF025377; AAC24981.1; -.
DR   HSSP; P23945; 1XUN.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 4.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    695       FOLLICLE STIMULATING HORMONE RECEPTOR.
FT   DOMAIN       18    366       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    367    387       1 (POTENTIAL).
FT   DOMAIN      388    398       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    399    421       2 (POTENTIAL).
FT   DOMAIN      422    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    444    465       3 (POTENTIAL).
FT   DOMAIN      466    485       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    486    508       4 (POTENTIAL).
FT   DOMAIN      509    528       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    529    550       5 (POTENTIAL).
FT   DOMAIN      551    573       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    574    597       6 (POTENTIAL).
FT   DOMAIN      598    608       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    609    630       7 (POTENTIAL).
FT   DOMAIN      631    695       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       44     68       LRR 1.
FT   REPEAT       69     93       LRR 2.
FT   REPEAT      119    143       LRR 3.
FT   REPEAT      170    192       LRR 4.
FT   REPEAT      193    216       LRR 5.
FT   REPEAT      218    240       LRR 6.
FT   DISULFID    442    517       BY SIMILARITY.
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT      2      2       S -> A (IN REF. 1).
FT   CONFLICT     13     13       T -> S (IN REF. 1).
FT   CONFLICT     60     60       V -> A (IN REF. 1).
FT   CONFLICT    166    166       V -> M (IN REF. 1).
FT   CONFLICT    215    215       Q -> H (IN REF. 1).
FT   CONFLICT    247    247       K -> R (IN REF. 1).
FT   CONFLICT    257    257       S -> T (IN REF. 1).
FT   CONFLICT    334    334       D -> N (IN REF. 1).
FT   CONFLICT    349    349       E -> K (IN REF. 1).
FT   CONFLICT    352    352       T -> A (IN REF. 1).
FT   CONFLICT    383    383       V -> E (IN REF. 1).
FT   CONFLICT    407    407       A -> T (IN REF. 1).
FT   CONFLICT    421    421       V -> I (IN REF. 1).
FT   CONFLICT    427    427       T -> S (IN REF. 1).
FT   CONFLICT    435    435       D -> N (IN REF. 1).
FT   CONFLICT    483    483       L -> V (IN REF. 1).
FT   CONFLICT    550    550       T -> I (IN REF. 1).
FT   CONFLICT    586    586       A -> V (IN REF. 1).
FT   CONFLICT    607    607       S -> L (IN REF. 1).
FT   CONFLICT    691    691       R -> H (IN REF. 1).
SQ   SEQUENCE   695 AA;  78172 MW;  E9EBEDB29C79C450 CRC64;

  Query Match             16.5%;  Score 136.5;  DB 1;  Length 695;
  Best Local Similarity   26.8%;  Pred. No. 3.4e-05;
  Matches   42;  Conservative   36;  Mismatches   62;  Indels   17;  Gaps    6;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| |:: | :|: ||::|   |  :: :|    | ::          :  :|||  ::|
Db        527 SQLYVVSL-LVLNVLAFVVICGCYTHIYLTVRNPNIMSS------SSDTKIAKRMAMLIF 579

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |   ||      |  : :  ::: :  |||  || || : |: | :
Db        580 TDFLCMAPISFFAISASLKVPLITVSKSKILLVLFYPINSCANPFLYAIFTKNFRRDVFI 639

Qy        116 MIHRFWHNYRQRKSMDSKGI-RKHMLHHSSGGKCGHC 151
              :: :|     | ::  :: :   | :|  :    |||
Db        640 LLSKFGCYEMQAQTYRTENLSTAHNIHPRN----GHC 672


RESULT 13
LSHR_MOUSE
ID   LSHR_MOUSE     STANDARD;      PRT;   700 AA.
AC   P30730;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR OR LHR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92165799; PubMed=1311310;
RA   Gudermann T., Birnbaumer M., Birnbaumer L.;
RT   "Evidence for dual coupling of the murine luteinizing hormone
RT   receptor to adenylyl cyclase and phosphoinositide breakdown and Ca2+
RT   mobilization. Studies with the cloned murine luteinizing hormone
RT   receptor expressed in L cells.";
RL   J. Biol. Chem. 267:4479-4488(1992).
RN   [2]
RP   SEQUENCE OF 1-58 FROM N.A.
RX   MEDLINE=93093308; PubMed=1459341;
RA   Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T., Sprengel R.;
RT   "The murine luteinizing hormone and follicle-stimulating hormone
RT   receptor genes: transcription initiation sites, putative promoter
RT   sequences and promoter activity.";
RL   Mol. Cell. Endocrinol. 88:55-66(1992).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
CC       THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
```

```
CC       ACTIVATE ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; S49753; AAB24402.1; -.
DR   EMBL; M81310; AAA39432.1; -.
DR   EMBL; M87571; AAA39433.1; -.
DR   PIR; A42395; A42395.
DR   HSSP; P22888; 1LUT.
DR   MGD; MGI:96783; Lhcgr.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    700       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                                RECEPTOR.
FT   DOMAIN       27    362       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    363    390       1 (POTENTIAL).
FT   DOMAIN      391    399       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    400    422       2 (POTENTIAL).
FT   DOMAIN      423    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    444    466       3 (POTENTIAL).
FT   DOMAIN      467    486       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    487    509       4 (POTENTIAL).
FT   DOMAIN      510    529       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    530    551       5 (POTENTIAL).
FT   DOMAIN      552    574       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    575    598       6 (POTENTIAL).
FT   DOMAIN      599    609       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    610    631       7 (POTENTIAL).
FT   DOMAIN      632    700       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       52     75       LRR 1.
FT   REPEAT      126    150       LRR 2.
FT   REPEAT      176    200       LRR 3.
FT   REPEAT      225    248       LRR 4.
FT   DISULFID    443    518       BY SIMILARITY.
FT   CARBOHYD    103    103       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    178    178       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    303    303       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    317    317       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   700 AA;  78214 MW;  8A6840A011E1E014 CRC64;

  Query Match             16.2%;  Score 134;  DB 1;  Length 700;
  Best Local Similarity   28.0%;  Pred. No. 5.9e-05;
  Matches   35;  Conservative   29;  Mismatches   49;  Indels   12;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
                :|:| ::| | :|  ||::|   |  ::::|    :||        |:  :||:   ::|
Db        528 SQVYILSILL-LNAVAFVVICACYVRIYFAVQNPELTAP------NKDTKIAKKMAILIF 580

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              ||  |  || | |   :     |  | :  ::: :  :||  || ||  : |:  |:
Db        581 TDFTCMAPISFFAISAAFKVPLITVTNSKVLLVLFYPVNSCANPFLYAVFTKAFQRDFFL 640
```

```
Qy        116 MIHRF 120
                :: ||
Db        641 LLSRF 645

RESULT 14
LSHR_RAT
ID   LSHR_RAT       STANDARD;      PRT;   700 AA.
AC   P16235; P70646; Q63807; Q63808; Q63809;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lutropin-choriogonadotropic hormone receptor precursor (LH/CG-R)
DE   (LSH-R) (Luteinizing hormone receptor).
GN   LHCGR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89332512; PubMed=2502842;
RA   McFarland K.C., Sprengel R., Phillips H.S., Koehler M.,
RA   Rosemblit N., Nikolics K., Segaloff D.L., Seeburg P.H.;
RT   "Lutropin-choriogonadotropin receptor: an unusual member of the G
RT   protein-coupled receptor family.";
RL   Science 245:494-499(1989).
RN   [2]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   STRAIN=Sprague-Dawley; TISSUE=Ovary;
RX   MEDLINE=92347604; PubMed=1353463;
RA   Aatsinki J.T., Pietila E.M., Lakkakorpi J.T., Rajaniemi H.J.;
RT   "Expression of the LH/CG receptor gene in rat ovarian tissue is
RT   regulated by an extensive alternative splicing of the primary
RT   transcript.";
RL   Mol. Cell. Endocrinol. 84:127-135(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91209270; PubMed=2019252;
RA   Koo Y.B., Slaughter R.G., Ji T.H.;
RT   "Structure of the luteinizing hormone receptor gene and multiple
RT   exons of the coding sequence.";
RL   Endocrinology 128:2297-2308(1991).
RN   [4]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=91006819; PubMed=1976554;
RA   Bernard M.P., Myers R.V., Moyle W.R.;
RT   "Cloning of rat lutropin (LH) receptor analogs lacking the soybean
RT   lectin domain.";
RL   Mol. Cell. Endocrinol. 71:R19-R23(1990).
RN   [5]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=91126285; PubMed=2281186;
RA   Segaloff D.L., Sprengel R., Nikolics K., Ascoli M.;
RT   "Structure of the lutropin/choriogonadotropin receptor.";
RL   Recent Prog. Horm. Res. 46:261-303(1990).
RN   [6]
RP   SEQUENCE OF 295-700 FROM N.A.
RX   MEDLINE=91060531; PubMed=2174034;
RA   Tsai-Morris C.H., Buczko E., Wang W., Dufau M.L.;
RT   "Intronic nature of the rat luteinizing hormone receptor gene defines
RT   a soluble receptor subspecies with hormone binding activity.";
RL   J. Biol. Chem. 265:19385-19388(1990).
RN   [7]
RP   SEQUENCE OF 27-37.
RX   MEDLINE=89174723; PubMed=2925659;
RA   Roche P.C., Ryan R.J.;
RT   "Purification, characterization, and amino-terminal sequence of rat
RT   ovarian receptor for luteinizing hormone/human choriogonadotropin.";
RL   J. Biol. Chem. 264:4636-4641(1989).
RN   [8]
RP   MUTAGENESIS.
```

```
RX   MEDLINE=91332007; PubMed=1714448;
RA   Ji I., Ji T.H.;
RT   "Asp383 in the second transmembrane domain of the lutropin receptor
RT   is important for high affinity hormone binding and cAMP production.";
RL   J. Biol. Chem. 266:14953-14957(1991).
CC   -!- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE. THE
CC       ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC       ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Some isoforms may
CC       be secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=11;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=P16235-1; Sequence=Displayed;
CC       Name=1759;
CC         IsoId=P16235-2; Sequence=VSP_001969, VSP_001977, VSP_001978;
CC       Name=1834;
CC         IsoId=P16235-3; Sequence=VSP_001977, VSP_001978;
CC       Name=1950;
CC         IsoId=P16235-4; Sequence=VSP_001968;
CC       Name=2075;
CC         IsoId=P16235-5; Sequence=VSP_001971, VSP_001973;
CC       Name=C1;
CC         IsoId=P16235-6; Sequence=VSP_001975, VSP_001976;
CC       Name=C2;
CC         IsoId=P16235-7; Sequence=VSP_001970;
CC       Name=EA2;
CC         IsoId=P16235-8; Sequence=VSP_001972;
CC       Name=EB;
CC         IsoId=P16235-9; Sequence=VSP_001972, VSP_001977, VSP_001978;
CC       Name=B1;
CC         IsoId=P16235-10; Sequence=VSP_001972, VSP_001974, VSP_001979;
CC       Name=B3;
CC         IsoId=P16235-11; Sequence=VSP_001974, VSP_001979;
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M26199; AAA41528.1; -.
DR   EMBL; M61212; AAA41527.1; -.
DR   EMBL; M61211; AAA41527.1; JOINED.
DR   EMBL; S40803; AAB22680.1; -.
DR   EMBL; S40787; AAB22680.1; JOINED.
DR   EMBL; S40903; AAB22680.1; JOINED.
DR   EMBL; S40904; AAB22680.1; JOINED.
DR   EMBL; S40905; AAB22680.1; JOINED.
DR   EMBL; S40907; AAB22680.1; JOINED.
DR   EMBL; S40909; AAB22680.1; JOINED.
DR   EMBL; S40918; AAB22680.1; JOINED.
DR   EMBL; S40920; AAB22680.1; JOINED.
DR   EMBL; S40795; AAB22680.1; JOINED.
DR   EMBL; S40798; AAB22680.1; JOINED.
DR   EMBL; S40795; AAB22681.1; -.
DR   EMBL; S40787; AAB22681.1; JOINED.
DR   EMBL; S40903; AAB22681.1; JOINED.
DR   EMBL; S40904; AAB22681.1; JOINED.
DR   EMBL; S40905; AAB22681.1; JOINED.
DR   EMBL; S40907; AAB22681.1; JOINED.
DR   EMBL; S40909; AAB22681.1; JOINED.
DR   EMBL; S40918; AAB22681.1; JOINED.
DR   EMBL; S40920; AAB22681.1; JOINED.
DR   EMBL; S40803; AAB22682.2; -.
DR   EMBL; S40787; AAB22682.2; JOINED.
DR   EMBL; S40903; AAB22682.2; JOINED.
DR   EMBL; S40907; AAB22682.2; JOINED.
DR   EMBL; S40909; AAB22682.2; JOINED.
DR   EMBL; S40918; AAB22682.2; JOINED.
DR   EMBL; S40920; AAB22682.2; JOINED.
DR   EMBL; S40795; AAB22682.2; JOINED.
DR   EMBL; S40798; AAB22682.2; JOINED.
DR   EMBL; S40803; AAB22683.1; -.
DR   EMBL; S40787; AAB22683.1; JOINED.
DR   EMBL; S40903; AAB22683.1; JOINED.
DR   EMBL; S40904; AAB22683.1; JOINED.
DR   EMBL; S40905; AAB22683.1; JOINED.
DR   EMBL; S40907; AAB22683.1; JOINED.
DR   EMBL; S40909; AAB22683.1; JOINED.
DR   EMBL; S40918; AAB22683.1; JOINED.
DR   EMBL; S40920; AAB22683.1; JOINED.
DR   EMBL; S40795; AAB22683.1; JOINED.
DR   EMBL; S40798; AAB22683.1; JOINED.
DR   EMBL; S40803; AAB22684.2; -.
DR   EMBL; S40787; AAB22684.2; JOINED.
DR   EMBL; S40903; AAB22684.2; JOINED.
DR   EMBL; S40904; AAB22684.2; JOINED.
DR   EMBL; S40905; AAB22684.2; JOINED.
DR   EMBL; S40909; AAB22684.2; JOINED.
DR   EMBL; S40918; AAB22684.2; JOINED.
DR   EMBL; S40920; AAB22684.2; JOINED.
DR   EMBL; S40795; AAB22684.2; JOINED.
DR   EMBL; S40798; AAB22684.2; JOINED.
DR   EMBL; M68928; AAA41529.1; -.
DR   EMBL; M68917; AAA41529.1; JOINED.
DR   EMBL; M68918; AAA41529.1; JOINED.
DR   EMBL; M68919; AAA41529.1; JOINED.
DR   EMBL; M68920; AAA41529.1; JOINED.
DR   EMBL; M68921; AAA41529.1; JOINED.
DR   EMBL; M68922; AAA41529.1; JOINED.
DR   EMBL; M68923; AAA41529.1; JOINED.
DR   EMBL; M68925; AAA41529.1; JOINED.
DR   EMBL; M68926; AAA41529.1; JOINED.
DR   EMBL; M68927; AAA41529.1; JOINED.
DR   PIR; A49744; A49744.
DR   HSSP; P22888; 1LUT.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 2.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT   SIGNAL        1     26
FT   CHAIN        27    700       LUTROPIN-CHORIOGONADOTROPIC HORMONE
FT                                RECEPTOR.
FT   DOMAIN       27    362       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    363    390       1 (POTENTIAL).
FT   DOMAIN      391    399       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    400    422       2 (POTENTIAL).
FT   DOMAIN      423    443       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    444    466       3 (POTENTIAL).
FT   DOMAIN      467    486       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    487    509       4 (POTENTIAL).

  Query Match             16.2%;  Score 134;  DB 1;  Length 700;
  Best Local Similarity   27.2%;  Pred. No. 5.9e-05;
  Matches   34;  Conservative   31;  Mismatches   48;  Indels   12;  Gaps    4;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| ::| : :|: ||::|   |  ::::|    :||        |:  :||:   ::|
Db        528 SQVYILSILI-LNVVAFVVICACYIRIYFAVQNPELTAP------NKDTKIAKKMAILIF 580

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              ||  |  || | |   :     |  | :  ::: :  :||  || || : |: |:
```

```
Db        581 TDFTCMAPISFFAISAAFKVPLITVTNSKILLVLFYPVNSCANPFLYAIFTKAFQRDFLL 640

Qy        116 MIHRF 120
              :: ||
Db        641 LLSRF 645


RESULT 15
FSHR_MOUSE
ID   FSHR_MOUSE     STANDARD;      PRT;   692 AA.
AC   P35378; Q9D4C2; Q9QWV8;
DT   01-JUN-1994 (Rel. 29, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE   receptor).
GN   FSHR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv; TISSUE=Testis;
RA   Tena-Sempere M., Manna P.R., Huhtaniemi I.T.;
RT   "Molecular cloning of the mouse follicle stimulating hormone receptor
RT   complementary deoxyribonucleic acid: functional expression of
RT   alternatively spliced variants and receptor inactivation by a C566T
RT   transition in exon 7 of the coding sequence.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE OF 1-51 FROM N.A.
RX   MEDLINE=93093308; PubMed=1459341;
RA   Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T.,
RA   Sprengel R.;
RT   "The murine luteinizing hormone and follicle-stimulating hormone
RT   receptor genes: transcription initiation sites, putative promoter
RT   sequences and promoter activity.";
RL   Mol. Cell. Endocrinol. 88:55-66(1992).
CC   -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC       OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC       ADENYLATE CYCLASE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC       FSH/LSH/TSH SUBFAMILY.
CC   -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF095642; AAC67559.1; -.
DR   EMBL; AK016635; BAB30351.1; -.
DR   EMBL; S49632; AAB24401.1; -.
DR   EMBL; M87570; AAA37641.1; -.
DR   MGD; MGI:95583; Fshr.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000372; LRR_Nterm.
DR   Pfam; PF00001; 7tm_1; 1.
DR   Pfam; PF00560; LRR; 4.
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   SMART; SM00013; LRRNT; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW   Phosphorylation; Repeat; Leucine-rich repeat.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18    692       FOLLICLE STIMULATING HORMONE RECEPTOR.
FT   DOMAIN       18    365       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    366    386       1 (POTENTIAL).
FT   DOMAIN      387    397       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    398    420       2 (POTENTIAL).
FT   DOMAIN      421    442       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    443    464       3 (POTENTIAL).
FT   DOMAIN      465    484       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    485    507       4 (POTENTIAL).
FT   DOMAIN      508    527       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    528    549       5 (POTENTIAL).
FT   DOMAIN      550    572       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    573    596       6 (POTENTIAL).
FT   DOMAIN      597    607       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    608    629       7 (POTENTIAL).
FT   DOMAIN      630    692       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       44     68       LRR 1.
FT   REPEAT       69     93       LRR 2.
FT   REPEAT      119    143       LRR 3.
FT   REPEAT      170    192       LRR 4.
FT   REPEAT      193    216       LRR 5.
FT   REPEAT      218    240       LRR 6.
FT   DISULFID    441    516       BY SIMILARITY.
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    436    436       Q -> K (IN REF. 2).
SQ   SEQUENCE   692 AA;  77769 MW;  4B57229180563A44 CRC64;

  Query Match             16.0%;  Score 132.5;  DB 1;  Length 692;
  Best Local Similarity   26.6%;  Pred. No. 8e-05;
  Matches   42;  Conservative   30;  Mismatches   57;  Indels   29;  Gaps    5;

Qy          1 AQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFFIVF 60
              :|:| :|: | :|  ||::|   |  :: :|    | ::         ::  :|||   ::|
Db        526 SQLYVMAL-LVLNALAFVVICGCYTHIYLTVRNPNIVSS------SRDTKIAKRMATLIF 578

Qy         61 TDALCWIPI-FVAKPLSLLQVEIPGTITSWVVIGYSAINSALNPILYTLTTRPFKE---- 115
              || ||  || | |   ||      |  :      ::: :  |||  || || : |: |:
Db        579 TDFLCMAPILFFAISASLKVPLITVSKAKILLVLFYPINSCANPFLYAIFTKNFRRDFFV 638

Qy        116 MIHRF----------------WHNYRQRKSMDSKGIR 136
              :: :|                ||:  ||:  |   |
Db        639 LMSKFGCYEVQAQIYKTETSSITHNFHSRKNPCSSAPR 676


Search completed: September 25, 2003, 14:59:53
Job time : 12 secs
```